

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2003, 06:36:30 ; Search time 4701 Seconds

(without alignments)
10883.372 Million cell updates/sec

Title: US-10-009-782-1

Perfect score: 1758

Sequence: 1 gaattccactgacgcgga.....ccctgagctacgagaacctt 1758

Scoring table: IDENTITY-MDC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Maximum Match 100%

Listing first 45 summaries

Database :

```

1:  gb_ba:*
2:  gb_hg:*
3:  gb_in:*
4:  gb_om:*
5:  gb_ov:*
6:  gb_pat:*
7:  gb_ph:*
8:  gb_pl:*
9:  gb_pr:*
10: gb_ro:*
11:  gb_sts:*
12:  gb_sy:*
13:  gb_un:*
14:  gb_vi:*
15:  em_ba:*
16:  em_fun:*
17:  em_hum:*
18:  em_in:*
19:  em_mu:*
20:  em_om:*
21:  em_or:*
22:  em_ov:*
23:  em_pat:*
24:  em_ph:*
25:  em_pl:*
26:  em_ro:*
27:  em_sts:*
28:  em_un:*
29:  em_vi:*
30:  em_hg_hum:*
31:  em_hg_inv:*
32:  em_hg_other:*
33:  em_hg_mus:*
34:  em_hg_pln:*
35:  em_hg_rtd:*
36:  em_hg_mam:*
37:  em_hg_vrt:*
38:  em_sy:*
39:  em_hggo_hum:*
40:  em_hggo_mus:*
41:  em_hggo_other:*

```

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1758	100.0	1758	6	BD011857	BD011857 Transform
2	1758	100.0	1758	6	E59056	E59056 Transformed
3	1758	100.0	1758	23	BD004254	BD004254 Transform
4	1739.4	98.9	1980	1	S80683	S80683 dan-D-amino
5	1453.4	82.7	1455	1	AFANADA	D45918 Alcaligenes
6	1088.6	61.9	1455	1	AF332548	AF332548 Alcaligen
7	693.6	39.5	1497	1	AF332548	D45919 Alcaligenes
8	533.6	30.4	208050	1	AL646083	AL646083 Ralstonia
9	455.8	25.9	1467	1	AF332548	D50061 Alcaligenes
10	306	17.4	110000	2	LMFCHR34_10	Continuation (11 o
11	171.8	9.8	22251	1	SC2K36	Continuation (11 o
12	116	6.6	82746	1	AF435501	AF435501 Actinosyn
13	107.4	6.1	28751	1	SC6A11	AL138978 Streptomy
14	99.2	5.6	14710	1	AE006004	AE006004 Caulobact
15	99.2	5.6	123580	6	AX211705	AX211705 Sequence
16	99.2	5.6	123580	6	AX211705	AF263912 Streptomy
17	99.2	5.6	123580	6	AX211705	AX211739 Sequence
18	97.2	5.5	113193	1	AF357202	AF357202 Streptomy
19	94.2	5.4	8438	14	HEPVIE	X15120 Pseudorabie
20	94.2	5.4	8438	14	HEPVIE	M57505 Pseudorabie
21	94	5.3	4039	1	AF425991	AF425991 Streptomy
22	94	5.3	34766	1	SCC121	AL137166 Streptomy
23	93.6	5.3	110000	2	LMFCHR32_06	Continuation (7 of
24	93.4	5.3	27541	6	AX211706	AX211706 Sequence
25	91.4	5.2	37750	1	SC6D7	AL133213 Streptomy
26	91.4	5.2	35033	1	SC130A	AL096811 Streptomy
27	91	5.2	35654	1	SC7H1	AL021411 Streptomy
28	89.8	5.1	15141	14	SHIPIRE	M34651 Pseudorabie
29	89.6	5.1	33779	1	SCGD3	AL096822 Streptomy
30	88.8	5.1	110000	2	LMFCHR36_31	Continuation (32 o
31	88.2	5.0	38404	1	SC2G5	AL035478 Streptomy
32	88	5.0	13848	1	AE005104	AE005104 Halobacte
33	87.8	5.0	1746	1	AF425992	AF425992 Streptomy
34	87.8	5.0	30657	1	SCD12A	AL357524 Streptomy
35	86.8	4.9	41906	1	SC5C7	AL031515 Streptomy
36	86	4.9	88421	6	AX417445	AX417445 Sequence
37	85.6	4.9	5520	3	AC084329	AC084329 Leishman
38	85.6	4.9	49736	1	AF319998	AF319998 Stigmatel
39	85.2	4.8	25362	1	SCF56	AL133424 Streptomy
40	84.6	4.8	109519	6	AX195929	AX195929 Sequence
41	84	4.8	1839	6	AX196078	AX196078 Sequence
42	83.8	4.8	3578	1	AF005277	AF005277 Cellulomo
43	83.8	4.8	3619	1	SCVALSFP	Y13070 S.coelicolo
44	83.8	4.8	24800	1	SC9B1	AL049727 Streptomy
45	83.8	4.8	110000	2	LMFCHR36_03	Continuation (4 of

ALIGNMENTS

```

RESULT 1
BD011857
LOCUS      BD011857
DEFINITION Transformant microorganism and process for producing
ACCESSION  BD011857
VERSION    BD011857.1 GI:22092046
KEYWORDS   D-aminocyclase.
SOURCE     unidentified.
ORGANISM   unidentified.
REFERENCE  1 (bases 1 to 1758)
AUTHORS   Takeuchi,K., Koide,Y., Hirose,Y., Moriyuchi,M. and Isobe,K.
TITLE      Transformant microorganism and process for producing D-aminocyclase
JOURNAL    Patent: WO 0078926-A 1 28-DEC-2000.

```

AMANO PHARMACEUTICAL CO LTD, KENICHI TAKEUCHI, YOSHINO KOIDE,
YOSHIIHIKO HIROSE, MITSUAKI MORIGUCHI, KIMIYASU ISOBE
OS Alcaligenes xylosoxydans subsp. xylosoxydans
PN MO 0078926-A/1
PD 28-DEC-2000
PF 15-JUN-2000 WO 2000JP003932
PI 17-JUN-1999 JP 99P 170555
PI KENICHI TAKEUCHI, YOSHINO KOIDE, YOSHIIHIKO HIROSE, PI
MITSUAKI MORIGUCHI,
PI KIMIYASU ISOBE
PC C12N1/21, C12N15/52, C12N9/80
FH

FEATURES
source Location/Qualifiers
1. 1758
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 281 a 662 c 568 g 247 t
ORIGIN

Query Match 100.0% Score 1758; DB 6; Length 1758;
Best Local Similarity 100.0%; Pred. No. 9e-183;
Matches 1758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCACCTGATCGGGAGAGAGATTTCATGTCCTCAATCCGATTCGAGCCCTTC 60
DB 1 GAATTCACCTGATCGGGAGAGAGATTTCATGTCCTCAATCCGATTCGAGCCCTTC 60
QY 61 GACCTGCTGCTCGGGGGGACCCCTCATCGAGGCAACACCCCGGGGGGCGCC 120
DB 61 GACCTGCTGCTCGGGGGGACCCCTCATCGAGGCAACACCCCGGGGGGCGCC 120
QY 121 GACCTGGGCGTGGCGGCGACCCGATCGCGCATCGGCGATCTGTGGAGCGCGCCG 180
DB 121 GACCTGGGCGTGGCGGCGACCCGATCGCGCATCGGCGATCTGTGGAGCGCGCCG 180
QY 181 CACACCGGGGTGAGCTGTGCGGGGCTGTGTGTCGCGCCGCTTCATGATCGACACAC 240
DB 181 CACACCGGGGTGAGCTGTGCGGGGCTGTGTGTCGCGCCGCTTCATGATCGACACAC 240
QY 241 CACGACGACACTTACCTCTCAGAGCGTGCAGACATGAGCGCCAGAGATCTCGAGGGCGTC 300
DB 241 CACGACGACACTTACCTCTCAGAGCGTGCAGACATGAGCGCCAGAGATCTCGAGGGCGTC 300
QY 301 ACCAGGCTGTACGAGGCAATTTGGGGCATCAAGCTTGGGGCGGCTTGGCGCAACCCG 360
DB 301 ACCAGGCTGTACGAGGCAATTTGGGGCATCAAGCTTGGGGCGGCTTGGCGCAACCCG 360
QY 361 CCGGCGCCCTGAGCTCTGTGAGCAAGGCGGCTTACCGCTTTCGAGCGCTTTCGCGAC 420
DB 361 CCGGCGCCCTGAGCTCTGTGAGCAAGGCGGCTTACCGCTTTCGAGCGCTTTCGCGAC 420
QY 421 TACCTGAGCGGCTTGGGCGCACGCGGCGGCGGCTCAACGCGCGCTGTATGGTGGCCAT 480
DB 421 TACCTGAGCGGCTTGGGCGCACGCGGCGGCGGCTCAACGCGCGCTGTATGGTGGCCAT 480
QY 481 TCAACGCTGCGGCGGCGGTATGCGGACCTTGGAGCGGCGCGCCACGAGAGAAATC 540
DB 481 TCAACGCTGCGGCGGCGGTATGCGGACCTTGGAGCGGCGCGCCACGAGAGAAATC 540
QY 541 GCGGCGATGCGGAGCTTGGCGGAGAGGCAATGGCGAGCGGCGGCTTTCGAGCC 600
DB 541 GCGGCGATGCGGAGCTTGGCGGAGAGGCAATGGCGAGCGGCGGCTTTCGAGCC 600
QY 601 GCGGCGCTTTCACCG 660
DB 601 GCGGCGCTTTCACCG 660
QY 661 CCGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 CCGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 721 GTGGCGCGCTGAGGAACTTCGCGATGGCGCGGAGCTGGAGCTGGCGGCTGTGATC 780

DB 721 GTGGCGCGCTGAGGAACTTCGCGATGGCGCGGAGCTGGAGCTGGCGGCTGTGATC 780
QY 781 TCGCACCAACAAGGTCATGGGCGGAGCCCAATTTGGCGGCTCGCGGAGAGAGCTGCGCGTG 840
DB 781 TCGCACCAACAAGGTCATGGGCGGAGCCCAATTTGGCGGCTCGCGGAGAGAGCTGCGCGTG 840
QY 841 ATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
DB 841 ATGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
QY 901 TCCACATGCTCAAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
DB 901 TCCACATGCTCAAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 960
QY 961 AAGCCCTTCCCGCAATGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
DB 961 AAGCCCTTCCCGCAATGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1020
QY 1021 TCCAGTACGACGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 1021 TCCAGTACGACGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
QY 1081 GAACCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 GAACCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 CCGGACGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
DB 1141 CCGGACGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1200
QY 1201 TATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
DB 1201 TATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1260
QY 1261 ACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
DB 1261 ACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1320
QY 1321 CTGGTGTGTTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
DB 1321 CTGGTGTGTTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1380
QY 1381 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
DB 1381 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1440
QY 1441 TTCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
DB 1441 TTCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1500
QY 1501 CCGTTACATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
DB 1501 CCGTTACATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1560
QY 1561 CGGTACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
DB 1561 CGGTACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1620
QY 1621 GTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1680
DB 1621 GTGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1680
QY 1681 TCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
DB 1681 TCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1740
QY 1741 CTGAGCTACGAGAGCTT 1758
DB 1741 CTGAGCTACGAGAGCTT 1758

RESULT 2

ES9056
LOCUS E59056 1758 bp DNA linear PAT 31-JAN-2002
DEFINITION Transformed microorganism and process for producing D-aminomylase.
ACCESSION E59056
VERSION E59056.1 GI:18629918
KEYWORDS JP 2001000185-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1758)
AUTHORS Takeuchi,K., Koide,Y., Hirose,Y., Moriguchi,M. and Isohe,K.
TITLE Transformed microorganism and process for producing D-aminomylase
JOURNAL Patent: JP 2001000185-A 1 09-JAN-2001;
AMANO PHARMACEUT CO LTD
OS Alcaligenes xylosoxydans subsp. xylosoxydans
PN JP 2001000185-A/1
PD 09-JAN-2001
PF 17-JUN-1999 JP 1999170555
PR
PI KENICHI TAKEUCHI,YOSHINAO KOIDE,YOSHIHIKO HIROSE, PI
MITSUAKI MORIGUCHI,
PI KIMIYASU ISOHE
PC C12N15/09,C12N1/21,C12N9/80/(C12N15/09,C12R1:05),C12N9/80,
PC C12R1:19),
PC C12N15/00,(C12N15/00,C12R1:05)
CC
CC
FH
FT
FT
KEY Key
FT source
Location/Qualifiers
1.1758
Location/Qualifiers
1.1758
/organism='Alcaligenes xylosoxydans subsp. xylosoxydans'
/db_xref='taxon:32644'
/organism='unidentified'
281 a 662 c 568 g 247 t

Query Match 100.0%; Score 1758; DB 6; Length 1758;
Best Local Similarity 100.0%; Pred. No. 9e-183;
Matches 1758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCACCTGATCGCGAAGAGAGATTTCATGTCCTCAATCCGATTCGAGCCCTTC 60
DB 1 GAATTCACCTGATCGCGAAGAGAGATTTCATGTCCTCAATCCGATTCGAGCCCTTC 60
QY 61 GACCTGCTGCGGGGGGACACCTTCATCGAGGCGACAAACCCCGGGGGCGCGCC 120
DB 61 GACCTGCTGCGGGGGGACACCTTCATCGAGGCGACAAACCCCGGGGGCGCGCC 120
QY 121 GACCTGCGCTGCGGGGACCGCATCGCCGATCGGCGATGTCGGAGCGCGCGCCG 180
DB 121 GACCTGCGCTGCGGGGACCGCATCGCCGATCGGCGATGTCGGAGCGCGCGCCG 180
QY 181 CACACCGGGGTCGACGTGTGCGGCTGTGTCGCGCCCGCTTCATGACATCGCACAC 240
DB 181 CACACCGGGGTCGACGTGTGCGGCTGTGTCGCGCCCGCTTCATGACATCGCACAC 240
QY 241 CAGCAGCAACAATACCTGCTAGGCGTTCGACATGACGCCCAAGATCTCGCAGGGCTC 300
DB 241 CAGCAGCAACAATACCTGCTAGGCGTTCGACATGACGCCCAAGATCTCGCAGGGCTC 300
QY 301 ACCAGGTGTGTCAGGGCAATTTGGGATCAGCTGAGGGGGCGGGGCGACAGCCACCCG 360
DB 301 ACCAGGTGTGTCAGGGCAATTTGGGATCAGCTGAGGGGGCGGGGCGACAGCCACCCG 360
QY 361 CCGCGCCCCCTGACCTGTGTGAGCAAGGCGCTTACCTTTTCGAGCGCTTTCGCCGAC 420
DB 361 CCGCGCCCCCTGACCTGTGTGAGCAAGGCGCTTACCTTTTCGAGCGCTTTCGCCGAC 420
QY 421 TACCTGAGCGGTGCGGGGCGACCGCGCGCGCTCAACGCCGCTGATGTGGGCGAT 480
DB 421 TACCTGAGCGGTGCGGGGCGACCGCGCGCGCTCAACGCCGCTGATGTGGGCGAT 480

QY 481 TCAGCGCTGCGCGCGCGGCTATGCCGACTTTCGACGCGCGCCGACGAGAGAAATC 540
DB 481 TCAGCGCTGCGCGCGCGGCTATGCCGACTTTCGACGCGCGCCGACGAGAGAAATC 540
QY 541 GCGGCGATCGGGGACCTGCGCGAGGAGGCAATGCGGCGCGGCGGCGGCGGCGGCGG 600
DB 541 GCGGCGATCGGGGACCTGCGCGAGGAGGCAATGCGGCGCGGCGGCGGCGGCGGCGG 600
QY 601 GCGGCGCTTTCACCG 660
DB 601 GCGGCGCTTTCACCG 660
QY 661 CCGCTGAGCGGCGGATGCGGCGATTCAGCGCCACCGCATGCGGCGAGGAGGAGGAG 720
DB 661 CCGCTGAGCGGCGGATGCGGCGATTCAGCGCCACCGCATGCGGCGAGGAGGAGGAG 720
QY 721 GTGGCGCGCGGTGAGGAAACCTTCGCGATGCGGCGCGCGCGCGCGCGCGCGCGCG 780
DB 721 GTGGCGCGCGGTGAGGAAACCTTCGCGATGCGGCGCGCGCGCGCGCGCGCGCGCG 780
QY 781 TCGCACACAAAGTATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 TCGCACACAAAGTATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 ATCGAGGCGCGCGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 ATCGAGGCGCGCGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 TCCACCATGCTCAGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
DB 901 TCCACCATGCTCAGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
QY 961 AAGCCCTTCCCGGAACCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 961 AAGCCCTTCCCGGAACCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
QY 1021 TCCAGTACGAGCTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 TCCAGTACGAGCTGTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 GAACCGGAGGTCGAGCGGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1081 GAACCGGAGGTCGAGCGGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 CCGCAGCAGAGCG 1200
DB 1141 CCGCAGCAGAGCG 1200
QY 1201 TATGCGCGGAGCTGTGCGCTTCCCGCTGAGAGCGCGGATGAGAGATGACCGGCGCTG 1260
DB 1201 TATGCGCGGAGCTGTGCGCTTCCCGCTGAGAGCGCGGATGAGAGATGACCGGCGCTG 1260
QY 1261 ACCGCGCGCGCTTCCCGCTGAGAGCGCGGATGAGAGATGACCGGCGCTG 1320
DB 1261 ACCGCGCGCGCTTCCCGCTGAGAGCGCGGATGAGAGATGACCGGCGCTG 1320
QY 1321 CTGCTGTGTTCGACCG 1380
DB 1321 CTGCTGTGTTCGACCG 1380
QY 1381 CCG 1440
DB 1381 CCG 1440
QY 1441 TTTACG 1500
DB 1441 TTTACG 1500
QY 1501 CCGTTACATTCGCGGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
DB 1501 CCGTTACATTCGCGGAGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
QY 1561 CGCTACATGCG 1620

Dd	1651	CGCTACATGGCCCTTCCTCCGCTTGCAGTAATAGGCCCAACCAGATATGTGGGCAAGAA	1620
Oy	1621	GTTATGGGCGCGCCCTTGCGGCGAGCGCAAAGCCCGAAAAATGACCTGCAAGACTTG	1680
Dd	1621	GTGAATGGGCGCGCCCTTGCGGCGAGCGCAAAGCCCGAAAAATGACCTGCAAGACTTG	1680
Oy	1681	TGCGAGGCCAGCGCATMGCGGTCTGACCCCTGTCCAAGGCCGAGCTGGGCCAGATCGCC	1740
Dd	1681	TGCGAGGCCAGCGCATMGCGGTCTGACCCCTGTCCAAGGCCGAGCTGGGCCAGATCGCC	1740
Oy	1741	CTGAGCTACGAGAACCTT 1758	
Dd	1741	CTGAGCTACGAGAACCTT 1758	
 RESULT 3 BD004254 standard; DNA: UNC; 1758 BP.			
ID	BD004254		
AC	BD004254;		
SV	BD004254.1		
DT	08-FEB-2002 (Rel. 70, Created)		
XX	08-FEB-2002 (Rel. 70, Last updated, Version 1)		
DE	Transformant microorganism and process for producing D-aminoacylase.		
XX	JP 03074329-T/1.		
KW			
XX	unidentified		
OS	unclassified.		
XX			
RN	[1]		
RP	1-1758		
RA	Tateuchi K., Koide Y., Hirose Y., Moriguchi M., Isobe K.;		
RT	"Transformant microorganism and process for producing D-aminoacylase";		
RL	Patent number JP03074329-T/1, 12-JAN-2001.		
XL	AMANO PHARMACEUTICAL CO LTD.		
CC	OS Alcaligenes xylosoxydans subsp. xylosoxydans		
CC	PN JP 03074329-T/1		
CC	PD 12-JAN-2001		
CC	PF 15-JUN-2000 JP 2000003932		
CC	PR 17-JUN-1999 JP 99P 170555		
CC	PI KENICHI TAKEUCHI, YOSHINAO KOIDE, YOSHIHIKO HIROSE,		
CC	MITSUAKI MORIGUCHI,		
CC	PI KIMIYASU ISOBE		
CC	PC C12N1/21,C12N15/52,C12N9/80		
CC	CC CC		
CC	FH key	Location/Qualifiers	
CC	FT source	1..1758	
CC	FT	/db_xref="taxon:32644"	
CC	FT	/organism="Alcaligenes xylosoxydans subsp. xylosoxydans"	
CC	XX		
FH	Key	Location/Qualifiers	
FT	source	1..1758	
FT		/db_xref="taxon:32644"	
FT		/organism="unidentified"	
XX			
SQ	Sequence 1758 BP; 281 A; 662 C; 568 G; 247 T; 0 other:		
 Query Match 100.0%; Score 1758; DB 23; Length 1758; Best Local Similarity 100.0%; Pred. No. 9e-183; Matches 1758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Oy	1	GAATTCCACTTATCGCGGAGAGAGATTTCCATGTCCCAATCCGATTCACGACCTTC	60
Dd	1	GAATTCCACTTATCGCGGAGAGAGATTTCCATGTCCCAATCCGATTCACGACCTTC	60
Oy	61	GACCTGCTCCTCGCGGCGGCAACCTCATCGACGGAGACAACCCCGGGGCGCGCC	120

Db	61	GACCTGTGCTCGCGGGCGGCGACCCCTCATCGAGGAGCAGCACCCCGGGGGCGCGCC	120
Oy	121	GACCTGGGGCGGGCGGCGGCGACCGCATCGCGCCCATCGGGGATCTGTCGGAGAGCGCGCGG	180
Db	121	GACCTGGGGCGGGCGGCGGCGACCGCATCGCGCCCATCGGGGATCTGTCGGAGAGCGCGCGG	180
Oy	181	CACACCGGGGTGAGCTGTCTGGGGCTGTGTGTGTGCGCCCGCTTATGAGCTCGCACAC	240
Db	181	CACACCGGGGTGAGCTGTCTGGGGCTGTGTGTGTGCGCCCGCTTATGAGCTCGCACAC	240
Oy	241	CACGACGACAACCTACCTGTCTAGGCGTCGCGACATGACGCCCAAGATCTCGCAGGGCGTC	300
Db	241	CACGACGACAACCTACCTGTCTAGGCGTCGCGACATGACGCCCAAGATCTCGCAGGGCGTC	300
Oy	301	ACGACGGGGGTACAGGGGCATTTGGGCAATCAGCTGTGGGCGCGCTGGCGACGCAACCCG	360
Db	301	ACGACGGGGGTACAGGGGCATTTGGGCAATCAGCTGTGGGCGCGCTGGCGACGCAACCCG	360
Oy	361	CCGCGCCCCCTTGAGACTGTGTGAGAGGAAGCGGCTTACCTGTTTGAGCGCTTCGCGAC	420
Db	361	CCGCGCCCCCTTGAGACTGTGTGAGAGGAAGCGGCTTACCTGTTTGAGCGCTTCGCGAC	420
Oy	421	TACCTGAGCGGCTGTGGGGGCCACGCGGGCGGCGTCAACGCGCGCTGTATGGTGGGCCAT	480
Db	421	TACCTGAGCGGCTGTGGGGGCCACGCGGGCGGCGTCAACGCGCGCTGTATGGTGGGCCAT	480
Oy	481	TCGACGCTCGCCCGCGGGGTATCGCGGATTCGAGACGGCGGCGCCACGAGAGGAATC	540
Db	481	TCGACGCTCGCCCGCGGGGTATCGCGGATTCGAGACGGCGGCGCCACGAGAGGAATC	540
Oy	541	CGGCGCATGCGGGGACCTGTGGCGAGGAAGCCATGGCCAGCGGCGCCATCGGCAATTCGACC	600
Db	541	CGGCGCATGCGGGGACCTGTGGCGAGGAAGCCATGGCCAGCGGCGCCATCGGCAATTCGACC	600
Oy	601	GCGCGCTTCTACCCGCGCGCGCGCGCCGCGCACCGGAGAGATCATCGAGGTGTGCGCG	660
Db	601	GCGCGCTTCTACCCGCGCGCGCGCGCGCCGCGCACCGGAGAGATCATCGAGGTGTGCGCG	660
Oy	661	CGGCTGAGCGGGCATTTGGCGGCATCTACGGCACCCATCGGAGAGGAGCGACATC	720
Db	661	CGGCTGAGCGGGCATTTGGCGGCATCTACGGCACCCATCGGAGAGGAGCGACATC	720
Oy	721	GTGGCGCGGTGAGGAACCTTCGCGCATGGCGCGGACCTGGAGCTGGCGGTGTGATC	780
Db	721	GTGGCGCGGTGAGGAACCTTCGCGCATGGCGCGGACCTGGAGCTGGCGGTGTGATC	780
Oy	781	TCGCAACCAAGGTATGTGGGCGACGCCAATTTTGCGCGCTCGCGGAGAGCGCTGCCGCTG	840
Db	781	TCGCAACCAAGGTATGTGGGCGACGCCAATTTTGCGCGCTCGCGGAGAGCGCTGCCGCTG	840
Oy	841	ATCGAGCGCGGCATATGGCGCGCGCGACAGAGCTGTGCTGTGACGCGTATTCCTACGTGGCGCG	900
Db	841	ATCGAGCGCGGCATATGGCGCGCGCGACAGAGCTGTGCTGTGACGCGTATTCCTACGTGGCGCG	900
Oy	901	TCGACATGTCTCAGGAGAGACCGGATCTGTGTGGCGGAGGACCATCATCTATCCTGTGGCTGC	960
Db	901	TCGACATGTCTCAGGAGAGACCGGATCTGTGTGGCGGAGGACCATCATCTATCCTGTGGCTGC	960
Oy	961	AAGCCCTTCCCCGAAGTGTAGCGGGGCGGACCTGGATGAAGTGTGCGGGCGAGGCGGGCAAA	1020
Db	961	AAGCCCTTCCCCGAAGTGTAGCGGGGCGGACCTGGATGAAGTGTGCGGGCGAGGCGGGCAAA	1020
Oy	1021	TCGCAAGTACAGCTGTGTGCCGAGACTGTCAAGCGGGCGGCATCTACTCATGTATGAGAC	1080
Db	1021	TCGCAAGTACAGCTGTGTGCCGAGACTGTCAAGCGGGCGGCATCTACTCATGTATGAGAC	1080
Oy	1081	GACCCGACGTGAGCGCATCTGTGGCGCTTGGGCGCGACCATGTATGCGGCTTCGAGCGGCTG	1140
Db	1081	GACCCGACGTGAGCGCATCTGTGGCGCTTGGGCGCGACCATGTATGCGGCTTCGAGCGGCTG	1140
Oy	1141	CCGACGACGAGCGCCGCGATTCGCGGCTGTGGGACCTTCCGCGGGTGTGTGGGCGAC	1200

Db 1141 CCGACGACGAGCGCCGCGATCCGCGCTGTGGGCGACCTTCCCGGGGTGTGGGCGAC 1200
 QY 1201 TATGCGCGGACCTGGGCGCTGTCCCGCTGGAGAGGGCGGTATGAGATGACCGGCGTC 1260
 Db 1201 TATGCGCGGACCTGGGCGCTGTCCCGCTGGAGAGGGCGGTATGAGATGACCGGCGTC 1260
 QY 1261 ACCGCGCGCGCTTGGCGCTGGGCGCGCGGCGACCTGCGAGCGCGGATCTTCCGCGAC 1320
 Db 1261 ACCGCGCGCGCTTGGCGCTGGGCGCGCGGCGACCTGCGAGCGCGGATCTTCCGCGAC 1320
 QY 1321 CTTGGTGTGTGGACCG 1380
 Db 1321 CTTGGTGTGTGGACCG 1380
 QY 1381 CCG 1440
 Db 1381 CCG 1440
 QY 1441 TTTACCG 1500
 Db 1441 TTTACCG 1500
 QY 1501 CCCTTACATCCG 1560
 Db 1501 CCCTTACATCCG 1560
 QY 1561 CGCTACATGCG 1620
 Db 1561 CGCTACATGCG 1620
 QY 1621 GTGATGCG 1680
 Db 1621 GTGATGCG 1680
 QY 1681 TCGCAGCG 1740
 Db 1681 TCGCAGCG 1740
 QY 1741 CTGAGCTACGAGAGCTT 1758
 Db 1741 CTGAGCTACGAGAGCTT 1758

RESULT 4

S80683

LOCUS

1980 bp DNA linear BCT 13-AUG-1996

DEFINITION

dan-D-aminocyclase [Alcaligenes xylosoxydans, ssp. xylosoxydans, A-6, Genomic, 1980 nt].

ACCESSION

S80683

VERSION

S80683.1

KEYWORDS

S80683.1 GI:1488356

SOURCE

ORGANISM

Achromobacter xylosoxydans A-6 ssp. xylosoxydans.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REMARK

FEATURES

source

gene

CDS

/note="D-aminocyclase, N-acyl-D-amino acid amidohydrolase"

201..1655

/gene="dan"

/note="3.5.1.14; Method: conceptual translation with partial peptide sequencing; This sequence comes from Fig. 3; N-acyl-D-amino acid amidohydrolase"

/transl_table=1

/product="D-aminocyclase"

/protein_id="AAB35881.1"

/db_xref="GI:1488357"

/translation="MSQSDQPFLLLAGGLTIDGNSNTPGRADLGYRGRIRIAISGLSDAAHPRVDVSGIVVAPGPTDSTHNDNLTIRRDMPKISQVTVTVVNGISILAPLANPPAPLDLDBSGSTFEFRADTDLRLTPAVNACAWGSHPLAAVAPLQRAATDEBETIAARLAEAMASGAIGSTGAFPPARATTEIIEVCRPLSAHGIIYATHMRDEBETIAARLAEAMASGAIGSTGAFPPARATTEIIEVCRPLSAHGIIYOVSLDAPYVAGSTMLKODRVLLAGRTIITWCKPELSELDDEVAERGSKYVVPLEPAGAIYFMDDEPVDORILAFGPTMISDGLPHDERHPRLMGTFRVIGHYARDLPLETETAVMKTGILAAFRGAGGLOAGYFADLVYFDPATVADTATFEHPTERAAHSHVYNAPVWQDAPFTGQAGRYLARTAA"

BASE COUNT

298 a 758 c 642 g 282 t

ORIGIN

Query Match 98.9%; Score 1739.4; DB 1; Length 1980;

Best Local Similarity 99.4%; Pred. No. 9.1e-181;

Matches 1746; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GAATTCACCTGATGCGGAGAGAGATTGATGTCCTCAATCGATTCGACCGCTTC 60
 Db 1 GAATTCACCTGATGCGGAGAGAGATTGATGTCCTCAATCGATTCGACCGCTTC 60
 QY 168 GATCGCCACTGTGATGCGCCCGCGGAGATTTCATTCGATTCGATTCGACCGCTTC 227
 Db 168 GATCGCCACTGTGATGCGCCCGCGGAGATTTCATTCGATTCGATTCGACCGCTTC 227
 QY 61 GACGTGCTGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 120
 Db 61 GACGTGCTGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 120
 QY 228 GACCTGCTGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 287
 Db 228 GACCTGCTGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 287
 QY 121 GACGTGCGGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 180
 Db 121 GACGTGCGGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 180
 QY 288 GACGTGCGGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 347
 Db 288 GACGTGCGGCTGCGGGGGGCGACCTCATCGAGCGGACACACACCGCGGGCGCGCC 347
 QY 181 CACACCCGGGTGACGTGTCGGGCTGTGTGCGCGCGCGCTTCATGATCGACAC 240
 Db 181 CACACCCGGGTGACGTGTCGGGCTGTGTGCGCGCGCGCTTCATGATCGACAC 240
 QY 348 CACACCCGGGTGACGTGTCGGGCTGTGTGCGCGCGCGCTTCATGATCGACAC 407
 Db 348 CACACCCGGGTGACGTGTCGGGCTGTGTGCGCGCGCGCTTCATGATCGACAC 407
 QY 241 CACGAGCAACTACCTGCTGAGCGGTGCGGACATGAGCGGACATGTCGACAGGCG 300
 Db 241 CACGAGCAACTACCTGCTGAGCGGTGCGGACATGAGCGGACATGTCGACAGGCG 300
 QY 408 CACGAGCAACTACCTGCTGAGCGGTGCGGACATGAGCGGACATGTCGACAGGCG 467
 Db 408 CACGAGCAACTACCTGCTGAGCGGTGCGGACATGAGCGGACATGTCGACAGGCG 467
 QY 301 ACCACGCTGCTGACGCGGCAATTGCGGATGACGCTGGCGCGCTGGCGACGCAACCG 360
 Db 301 ACCACGCTGCTGACGCGGCAATTGCGGATGACGCTGGCGCGCTGGCGACGCAACCG 360
 QY 468 ACCACGCTGCTGACGCGGCAATTGCGGATGACGCTGGCGCGCTGGCGACGCAACCG 527
 Db 468 ACCACGCTGCTGACGCGGCAATTGCGGATGACGCTGGCGCGCTGGCGACGCAACCG 527
 QY 361 CCGCGCCCTGCGACCTGCTGAGCAGAGCGGCTTACCGTTTGCAGCGCTGCCGAC 420
 Db 361 CCGCGCCCTGCGACCTGCTGAGCAGAGCGGCTTACCGTTTGCAGCGCTGCCGAC 420
 QY 528 CCGCGCCCTGCGACCTGCTGAGCAGAGCGGCTTACCGTTTGCAGCGCTGCCGAC 587
 Db 528 CCGCGCCCTGCGACCTGCTGAGCAGAGCGGCTTACCGTTTGCAGCGCTGCCGAC 587
 QY 421 TACCTGAGCGGCTTGGCGGCGACCGCGGCGCGCGCTGATGATGATGATGATGAT 480
 Db 421 TACCTGAGCGGCTTGGCGGCGACCGCGGCGCGCGCTGATGATGATGATGATGAT 480
 QY 588 TACCTGAGCGGCTTGGCGGCGACCGCGGCGCGCGCTGATGATGATGATGATGAT 647
 Db 588 TACCTGAGCGGCTTGGCGGCGACCGCGGCGCGCGCTGATGATGATGATGATGAT 647
 QY 481 TCAACGCTGCGCGCGCGCGGCTGATGCGGATGACGCGCGCGCGCGCGCGCGCGCG 540
 Db 481 TCAACGCTGCGCGCGCGCGGCTGATGCGGATGACGCGCGCGCGCGCGCGCGCGCG 540
 QY 648 TCAACGCTGCGCGCGCGGCTGATGCGGATGACGCGCGCGCGCGCGCGCGCGCG 707
 Db 648 TCAACGCTGCGCGCGCGGCTGATGCGGATGACGCGCGCGCGCGCGCGCGCGCG 707
 QY 541 GCGGCGATGCGGCGGCTGCGGAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 600
 Db 541 GCGGCGATGCGGCGGCTGCGGAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 708 GCGGCGATGCGGCGGCTGCGGAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 767
 Db 708 GCGGCGATGCGGCGGCTGCGGAGAGCGGCGGCGCGCGCGCGCGCGCGCGCGCG 767
 QY 601 GCGGCGCTTACCG 660
 Db 601 GCGGCGCTTACCG 660
 QY 768 GCGGCGCTTACCG 827
 Db 768 GCGGCGCTTACCG 827
 QY 661 CCGCTGAGCG 720
 Db 661 CCGCTGAGCG 720

D	b	61	GGAGCAAAACCCCGGGGGGGGGCGCCGACACTGTGGGCGTGGCGGAGACGCATGCCGCC	120
O	y	154	ATCGGCATCTGTGAGACCGCCCGCGACACCCGGGTGACGTGTGGGCTTGGTGGTC	213
D	b	121	ATCGGCATGTGTGAGACCGCCCGCGACACCCGGGTGACGTGTGGGCTTGGTGGTC	180
O	y	214	GGCGCCGGCTTCATGCACTGGCAACCCAGACAACTACTGTCTCAAGGCTGGCAGC	273
D	b	181	GGCGCCGGCTTCATGCACTGGCAACCCAGACAACTACTGTCTCAAGGCTGGCAGC	240
O	y	274	ATGACGCCCAAGATCTGCGACGGGCTACACAGGTGTACAGGGCAATTGGCGCATACG	333
D	b	241	ATGACGCCCAAGATCTGCGACGGGCTACACAGGTGTACAGGGCAATTGGCGCATACG	300
O	y	334	CTGGCGCCGCTGGCGCACGCCAACCCGCCGCCCTGTGACCTGCTGGACAGAGCGGC	393
D	b	301	CTGGCGCCGCTGGCGCACGCCAACCCGCCGCCCTGTGACCTGCTGGACAGAGCGGC	360
O	y	394	TCCTTACGTTTCGAGGCTTCGCCGACTACCTGTGAGACGCTTGGGGCCACCGCGGCC	453
D	b	361	TCCTTACGTTTCGAGGCTTCGCCGACTACCTGTGAGACGCTTGGGGCCACCGCGGCC	420
O	y	454	GTCACACCCGCTTATATGATGGGCAATTCAACGCTGGCGCGCGGTCTATGGCGGATTTG	513
D	b	421	GTCACACCCGCTTCAATGATGGGCAATTCAACGCTGGCGCGCGGTCTATGGCGGATTTG	480
O	y	514	CAGCGCGCCGCGCACGACGAGAAATTCGGGGCCATCGGAGCTGTGACGAGAACCATG	573
D	b	481	CAGCGCGCCGCGCACGACGAGAAATTCGGGGCCATCGGAGCTGTGAGAACCATG	540
O	y	574	GGCAGCGCGCGCATCGGCAATTTGACCGGGGCTTCTACCGCGCGCGCGCGCCAC	633
D	b	541	GGCAGCGCGCGCATCGGCAATTTGACCGGGGCTTCTACCGCGCGCGCGCGCCAC	600
O	y	634	ACCCAGAGATCATCGAGGTGTCCGGCGCTGAGGGCCATGAGGGCATCTAGCCAC	693
D	b	601	ACCCAGAGATCATCGAGGTGTCCGGCGCTGAGGGCCATGAGGGCATCTAGCCAC	660
O	y	694	CACATGCGCGACGAGAGGCGACACATCTGGCGCGCTGTGAGAGAACTTTCCTCCATCGC	753
D	b	661	CACATGCGCGCGACGAGAGGCGACACATCTGGCGCGCTGTGAGAGAACTTTCCTCCATCGC	720
O	y	754	CGCGAGCTGACATCGCGGGGTGTGATCTGCGACACAAAGTCAATGGGCCAGCCCAATTTC	813
D	b	721	CGCGAGCTGACATCGCGGGGTGTGATCTGCGACACAAAGTCAATGGGCCAGCCCAATTTC	780
O	y	814	GGCGCTCGCGCGAGAGCGTGGCGGTGATGAGAGCGCCATGGCGCGCAGACCTCG	873
D	b	781	GGCGCTCGCGCGAGAGCGTGGCGGTGATGAGAGCGCCATGGCGCGCAGACCTCG	840
O	y	874	CTGAGACGATCTCCTTACGTGTGGCGGCTCCACCANTCTCAACAGAGCGCTGTGTGTG	933
D	b	841	CTGAGACGATCTCCTTACGTGTGGCGGCTCCACCANTCTCAACAGAGCGCTGTGTGTG	900
O	y	934	GCCGAGCGCACATCATACCTGTGTGAAGCCCTTCCGAACTAGCGGGCGCACCTG	993
D	b	901	GCCGAGCGCACATCATACCTGTGTGAAGCCCTTCCGAACTAGCGGGCGCACCTG	960
O	y	994	GATGAGTGGGGGGCGCGCGCGCAATCAGTACGAGTGTGTCCGAGCTGTGACCG	1053
D	b	961	GATGAGTGGGGGGCGCGCGCGCAATCAGTACGAGTGTGTCCGAGCTGTGACCG	1020
O	y	1054	GCGCGCGCATCTACTTCATGATGAGAGAACCCGAGCTGACAGCATCTGTGGCTTGGC	1113
D	b	1021	GCGCGCGCATCTACTTCATGATGAGAGAACCCGAGCTGACAGCATCTGTGGCTTGGC	1080
O	y	1114	CCGACCATGATGGCTTCGACGGCTGTGCGGACGACGAGCGCCCGCATCGCGCTGTGG	1173
D	b	1081	CCGACCATGATGGCTTCGACGGCTGTGCGGACGACGAGCGCCCGCATCGCGCTGTGG	1140
O	y	1174	GGACACTTCCCGGGGGGTGAGGGGCACTATGCGCGGAGACTGTGGGCTGTTCCTGGAG	1233
D	b	1141	GGACACTTCCCGGGGGGTGAGGGGCACTATGCGCGGAGACTGTGGGCTGTTCCTGGAG	1200

QY	1234	ACGGGCGGTATGGAAGATGACC	GCGCCTGCACCCGCCCGCCTTGGCCTTGCGCGCGGCGG	1293
Db	1201	ACGGGCGGTATGGAAGATGACC	GCGCCTGCACCCGCCCGCCTTGGCCTTGCGCGCGGCGG	1260
QY	1294	CAGCTGCAGAGCGCGGTACTTC	TGCGCAACCTGGGTGGTTTGACCCCGGCAAGGTGGCCGAT	1353
Db	1261	CAGCTGCAGAGCGCGGTACTTC	TGCGCAACCTGGGTGGTTTGACCCCGGCAAGGTGGCCGAT	1320
QY	1354	ACCGCCACCTTGCAGAACACCT	TACCGAGCGCGCGCGCGCATCATTCGTTGACGTCAAC	1413
Db	1321	ACCGCCACCTTGCAGAACACCT	TACCGAGCGCGCGCGCGCATCATTCGTTGACGTCAAC	1380
QY	1414	GGCGGCGCGGTGTGGCAGAGAG	CAGCGCGTTACACCGGCGCAGATCCCGCGCGCTGTGCGA	1473
Db	1381	GGCGGCGCGGTGTGGCAGAGAG	CAGCGCGTTACACCGGCGCAGATCCCGCGCGCTGTGCGA	1440
QY	1474	CGCACGCGCGCGCTGA	1488	
Db	1441	CGCACGCGCGCGCTGA	1455	
RESULT 6	AF332548		1455 bp DNA linear BCT 04-MAR-2001	
LOCUS	Alcaligenes faecalis D-aminoacylase gene, complete cds.			
DEFINITION	AF332548			
ACCESSION	AF332548.1	GI:3194735		
VERSION				
KEYWORDS	Alcaligenes faecalis.			
SOURCE	Alcaligenes faecalis			
ORGANISM	Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;			
REFERENCE	1 (bases 1 to 1455)			
AUTHORS	Hsu,C.S., Lai,W.L., Chang,W.W., Yang,Y.B. and Tsai,Y.C.			
TITLE	D-aminoacylase			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 1455)			
AUTHORS	Hsu,C.S., Lai,W.L., Chang,W.W., Yang,Y.B. and Tsai,Y.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (28-Dec-2000) Institute of Biochemistry, National Yang-Ming University, 115, Li-Nong St. Sec. 2, Shih-Pai, Taipei, Taiwan 11221, R.O. China			
FEATURES	Location/Qualifiers			
Source	1..1455			
	/organism="Alcaligenes faecalis"			
	/db_xref="taxon:511"			
CDS	1..1455			
	/EC_number="3.5.1.81"			
	/note="Daap"			
	/codon_start=1			
	/transl_table=1			
	/product="D-aminoacylase"			
	/protein_id="AAK15530.1"			
	/db_xref="GI:3194736"			
	/translation="MSOPDATPPDYILSGGVIDGTNAPGLADVGRDIRIAVGDII SASSARRIDIVAGVVSFGFTDSHTHDNYLTKRRDMTPKISQVTTVYVNGCISILA PLAANPRLPIDLDLVGSSFFRARSFDLELRAPAIPAVNAACWVGSTLEAAVPDI BREYTADEIQMOALADALASGAIIGSTGAFPAAHASTEEIIEVCRLITGGVYL ATHNRDEGEHIVQALEETFRIGRELDVNVSHHKVMGKLNGRSKEITALIEAAMS ODVSLDAPIVYAGSTMCKODRVLLAGRTLLTWCKDPYELSPDRDLIELTAERGSKYDV VPEIQPRAGAIYFMNDPEPVQIILAFPTMISDCPLPHDERHPRLMGTFEPVLVHSR DLGLEPTEAYAKMTGILTAAEFLAEKRQVDPGYVDLVFDPATVADSATEHPETERE AAGHSYVNNAAMVEDSFSGHAGRVLNAGA"			
BASE COUNT	247 a 537 c 446 g 225 t			
ORIGIN				
Query Match:	61.9%	Score 1088.6;	DB 1;	Length 1455;
Best Local Similarity	84.3%;	Fold No. 6.6e-110;		
Matches 1226;	Conservative	0;	Mismatch 229;	Indels 0;
Gap 0;				
QY	34	ANGTCCCAATCCGATTCCAGCCTT	GCAGCGCTGCTGCGCGGCGGCAACCTCATCGAC	93


```

/product="N-acyl-D-amino acid amidohydrolase"
/protein_id="BA008350.1"
/db_xref="GI:1800111"
/translation="MTDSTLDDAQAQADFTIAGATLIDGGGPARQDGLAVGRIV
ALGDPHAPGVVIDARGLALAPFTIDSHDDGLLAHPELIPVSGITTVVGN
GISTAPLSRQIPQDLDELPELRFATFMDLRALAEPLVAVIPLVHTLRLVA
VDDTGRATADRAAMRALIDELAQAGFVSTGYPPASAPATDEIDVOPLRG
RAGDIYATHLRDEADHIVPAMEALIGRELDCRYVFSHHKLAGERNRISRETLDM
SRAAATOVICDCHRPATSTMLRLDRLASRLITMSKCPREATGDSRYMAELG
LDDEAIALAPAGAIYFLMDQADVNRIFSHLTVVSDGLPDPHPRMGITNV
LRTNVRDRLSLETAIRKMTGLAAQVGLERGLRKGIVHADLVLPFANTDTATF
SAPIQVSGIAHVNWNGROYMDERTGAERPOQVLAIPDIAIPWSQOSE"
BASE COUNT      235 a      592 c      467 g      203 t
ORIGIN

```

```

Query Match      39.5% Score 693.6 DB 1: Length 1497;
Best Local Similarity 67.7% Pred. No. 6,3e-67;
Matches 988; Conservative 0; Mismatches 469; Indels 3; Gaps 1;

QY 45 CGATTCCAGCCCTTCGACCTGCTCCGCGGCGGACCCCTCATGAGCGGACGAC 104
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 27 CGGTCCCGGCAAGCCGATTCATCATCCCGCGCCACCTCATGAGCGGCGGCGG 86
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 105 CCGCGGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 164
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 87 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 146
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 165 GTCGAGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 224
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 CGCCCATGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 206
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 225 CATGCACTGCAACCCAGCAGCAGCACTACTGCTGACGCGTCCGAGATGAGCCCAA 284
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 CATGCACTGCAACCCAGCAGCAGCACTACTGCTGACGCGTCCGAGATGAGCCCAA 266
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 285 GATCTGCGAGGCGGTACCAAGGTGTGTCAGGCGCAATTCGGGATCAGCTTGGCGCT 344
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 GGTCTGCGAGGCGGTACCAAGGTGTGTCAGGCGCAATTCGGGATCAGCTTGGCGCT 326
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 345 GCGGCAAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 404
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 GTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 386
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 405 CGAGCGCTTCCGCGGCTGACCTGAGCGGCTTGGCGGCGGCGGCGGCGGCGGCT 464
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 CGCAGCTTCCGCGGCTGAGCGGCTTGGCGGCGGCGGCGGCGGCGGCGGCT 446
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 465 CTGTATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 524
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 447 CCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 506
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 525 CACCGAGGAGAAATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 584
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 507 CACCGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 566
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 585 CATGCGGATTTGACGCGGCGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCT 644
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 567 CTTCGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 626
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 645 CATGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 701
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 627 CATGCGGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 686
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 702 CGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 761
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 687 CGACGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 746
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 762 GACGCTGCGGCTGATCTGACACCAAGGTCAATGGCGGCGGCGGCGGCT 821
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 747 GACGCTGCGGCTGATCTGACACCAAGGTCAATGGCGGCGGCGGCGGCT 806
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 822 GCGCGAGGCGGCTGATCTGAGGCGGCGGCGGCGGCGGCGGCGGCT 881
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 807 CCGCGAAGCGTGGACATGATCAGCGCGCGCGCGCGCGCGCGCGCGGCTGCTGACTG 866
QY 882 GATATCCCTAGTGGCGGCTCCACCATGCTCAGAGAGACCGGTGCTGGCGGACG 941
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 867 CCACCCCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 926
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 942 CACCATCATCAGCTGGTGGCAAGCCCTTCCCGGCGGCGGCGGCGGCGGCGGCT 1001
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 927 CACCGCTGATCAGCTGGTGGCAAGCCCTTCCCGGCGGCGGCGGCGGCGGCGGCT 986
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1002 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1061
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 987 CATGCGGCAAGTGGCGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1046
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1062 CATCTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1121
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1047 CATCTACTTATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1122 CATGCGGCTGCGAGGCGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1181
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1107 GGTGCGGCTGCGAGGCGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1166
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1182 CCGCGGCGGCTGCGAGGCGGCTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1241
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1167 CACCAAGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1226
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1242 ATGGAAGATGAGCGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCGGCTGCA 1301
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1227 CCACAAGATGAGCGGCGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1286
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1302 GCGCGGCTGCTGCGGCGGCGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCT 1361
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1287 GCGAGGCTGCTGCGGCGGCGGCTTACCGCGGCGGCGGCGGCGGCGGCGGCT 1346
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1362 CTTCGAAACCCCTTACCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1421
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1347 TTTTTCAGCTTCCCATTCAGTACAGTACAGGAGGCGGCGGCGGCGGCGGCT 1406
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1422 GGTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1407 GGTCTGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1466
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1482 CCGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 1501
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1467 TGCATTAACCTTGGAGTCAAC 1486
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 8
AL646083
LOCUS
DEFINITION
  Ralstonia solanacearum GMI1000 megaplasmid, complete sequence;
ACCESSION
  AL646083 AL646053
VERSION
  AL646083.1 GI:17431588
SOURCE
  Ralstonia solanacearum.
  Ralstonia solanacearum.
  Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
  Ralstonia.
REFERENCE
  1 (bases 1 to 208050)
  Salanoubat,M., Genin,S., Artiguenave,F., Gouzy,J., Mangenot,S.,
  Ariat,M., Billault,A., Brotier,P., Camus,J.C., Catolico,L.,
  Chandler,M., Choisme,N., Claudel-Renard,C., Cunne,S., Demange,N.,
  Gaspin,C., Lavie,M., Moisan,A., Robert,C., Saurin,M., Schlex,T.,
  Siguler,P., Thebaud,P., Whalen,M., Winkler,P., Levy,M.,
  Weissenbach,J. and Boucher,C.A.
  Genome sequence of the plant pathogen Ralstonia solanacearum
  Unpublished
  2 (bases 1 to 208050)
  Boucher,C.A.
  Direct Submission
  Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
  Creteil, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie

```

Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS, BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean Dausset-CEPH, 27 rue Juliette Dodu, 75010 Paris, France, LMGM CNRS 118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA UGV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27, F31326 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire INRA, BP27, F31326 Castanet-Tolosan Cedex
 Christian.Boucher@toulouse.inra.fr
 http://sequence.toulouse.inra.fr/R.solanacearum.html.

FEATURES
 source
 1. 208050
 /organism="Ralstonia solanacearum"
 /strain="GM11000"
 /db_xref="taxon:305"
 /plasmid="megaplasmid"
 135..947
 /gene="Rsp1115"
 /note="RS05505"
 135..947
 /gene="Rsp1115"
 /function="miscellaneous: not classified regulator"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Homology
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PUTATIVE TRANSCRIPTION REGULATOR PROTEIN"
 /protein_id="CAD18266.1"
 /db_xref="GI:17431589"
 /translation="MADSVYOQVADVPGVLVLSGRFSQSFERRHFLDPHVLGVTAG
 VORRFKGVYLVGPGCISLMPGEIHGVAEAGAYTIKFRLSOALLNADISG
 ETRPERAATLEDPALAGHLRLHDMOHSGASTLOVOSPLALLAFLIKOSALLP
 ETRGALSAPQOMORVDYCFRAHAGERITLDDLASLCALGRQFKOKQYTGAPPHAM
 LVRLRLQVCLSSRSKALAEVAGVEGFTDOSHFRNAPROAFGVPASHYS"
 1162..2259
 /gene="Rsp1116"
 /note="RS05504"
 1162..2259
 /gene="Rsp1116"
 /EC_number="4.1.2.5"
 /function="small molecule metabolism; degradation; amino
 acids"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PUTATIVE LOW SPECIFICITY L-THREONINE ALDOLASE
 PROTEIN"
 /protein_id="CAD18267.1"
 /db_xref="GI:17431590"
 /translation="MNRDPMPTTRPALGFTSDNIEGASPEVLAILAANGAAPYG
 ADCTARVEORTLSLFEFEKVPFLPPTGTANALCIGMPMPGNAVYCHPSHINDG
 CGAPEFTNGAKLVAVDGPSPAKIDYEKIRATRYKGVHSTOPACVSLTATEYSEL
 YTLNEITAIIGQLCKDARKIKLOMDSREFANALVSGSPAAATWAGVVALSFGATKNG
 VVAEAIVLFDLSLAEMGYRRKRAGHLFSKMRFLSAQIEAYLANDLMLRNARANDM
 ALRLARGLEGHAGSIVLGATEANIVCFRLPSTLLENLLQAGFRFFYHNMEPVVREVT
 SEATYADVDVNLHRMKAAAGPSADTATA"
 2257..5019
 /note="ACUR or RS05503
 predicted by Framed"
 /evidence=not_experimental
 complement(2462..3139)
 /gene="Rsp1117"
 /note="RS05502"
 complement(2462..3139)
 CDS

gene
 /gene="Rsp1117"
 /function="miscellaneous: unknown"
 /note="Product confidence : putative
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="PUTATIVE TRANSMEMBRANE PROTEIN"
 /protein_id="CAD18268.1"
 /db_xref="GI:17431591"
 /translation="MNDASLHETIISCFLEHORPPVYRIAEFRNDEAQNVALRAL
 ADNHGVLLPHPSDEVIAIHPFSAPPTCVKSGNRKMWNCAMCALGLVHLGASATI
 EFRIGAIIDHVVIIIRIENGRLDLDYVHPPIPKHAMDVITYCSVQLFRDAVDYD
 WCASRGICPDVPRPIEQVNNFAVEWYARRADADWTKMSVREAIELEFARHLSGPVWTI
 AAEARF"
 complement(3188..3595)
 /gene="Rsp1118"
 /note="RS05501"
 complement(3188..3595)
 /gene="Rsp1118"
 /function="miscellaneous: hypothetical/global homology"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 /protein_id="CAD18269.1"
 /db_xref="GI:17431592"
 /translation="MFDTKIAYVRDLATWOKLNTYAFILMSGITGASPEINGEAYRD
 AADNYTHSLCQGVIVISGDEFTLKAHRALEKRPKHGIYEDMSTGHDAANRAVF
 AEHPGDSARVYGLALREKTIYDKTGARMHP"
 complement(3673..4161)
 /gene="Rsp1119"
 /note="RS05500"
 complement(3673..4161)
 /gene="Rsp1119"
 /function="miscellaneous: hypothetical/global homology"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 /protein_id="CAD18270.1"
 /db_xref="GI:17431593"
 /translation="MGRSEYGETVCPTARTLAVYGDWYVLIIRLEFLGYKREERQA
 OTGMSHLSTRLKLEADGIVVAKHLSDBPPRHRKLSKGLDFTPLLSLSWSEK
 WGFKAKTAPALITTHROCGHETGLKLVCPACDEPGPKDATVTTLGASFKAERARD
 GA"
 4228..4725
 /gene="Rsp1120"
 /note="RS05499"
 4228..4725
 /gene="Rsp1120"
 /function="miscellaneous: hypothetical/global homology"
 /note="Product confidence : hypothetical
 Gene name confidence : hypothetical
 predicted by Codon_usage
 predicted by Homology
 predicted by Framed"
 /codon_start=1
 /evidence=not_experimental
 /transl_table=11
 /product="CONSERVED HYPOTHETICAL PROTEIN"
 CDS

TITLE	Studies on structure and function of bacterial N-acyl-D-amino acids amidohydrolase
JOURNAL	Thesis (1995) _FACULTY_/Applied Chemistry/_MAJOR_, Oita University
REFERENCE	2 (bases 1 to 1467)
AUTHORS	Makayama,M.
TITLE	Journal
source	Direct Submission Submitted (06-MAR-1995) Memoru Makayama, Oita University, Applied Chemistry; Dannoharu, Oita, Oita 870-11, Japan (Tel.:0975-69-3311(ex.746), Fax:0975-69-7957) Location/Qualifiers 1. 1467 /organism="Achromobacter xylosoxidans" /strain="A-6" /db_xref="taxon:85698"
CDS	1. 1467 /codon_start=1 /transl_table=11 /product="N-acyl-D-amino acids amidohydrolase" /protein_id="BA08778.1" /db_xref="GI:1707291" /translation="MOEKLDIVIEGWNIDLGPPRRRADVIRGERIAITGDSLAP ADRLDAGGRVARGPIIDTHGDHMDIEMVEKGLERMKTSOGITSVWNCISGA PGRNALALILDGSPLEFRDAMTYFGALEDRPHINVALYVAHLRIANMDDPAOPS AKDEKAHERMLADLLEANGVESTGLYLQPEGVAEQLDELAVLAARLRHSHTIR NEGDAEAAVDLVLAEGVRRTGCATLSHHCKIMBPAANNKSAATLANIDRAAGDVYA LDIPYPGSSSTILPERADQIDDIRITWSPTPECGSLIEIARNICSDVTAAARI CPAATIFAMDENEVERIRFOHECCWMSDGLPNANHPRPLMGSFTRYIGRVREAE LTLEAAVKMTALPARVYGDLRGRLVAGAADVVYDVCBRATWDATFLASAGI EHVLNCAVFPPQAPPSIRREIRLLIRDSINAGEFSR"
BASE COUNT	213 a 530 c 530 g 194 t
ORIGIN	
Query Match	25.9%; Score 455.8; DB 1; Length 1467;
Best Local Similarity	59.3%; Pred. No. 4.8e-41;
Matches	813; Conservative 0; Mismatches 552; Indels 6; Gaps 2;
Oy	59 TCGACTCTGCTGCTCGGGGGCGGACACCCTCATCGACGGCAGCAACCCCGGGCGGGCGG 118
Db	14 TGGATCTGTGTCATCGAAAGCGGCTGGTGATCGACGGCGCTGGCGGGCGCGCGCGG 73
Oy	119 CCGACTTGGGGGTGGGGGGCGGCGCACGCCATCCCGGCATCGGCATCTGTGCGAGCGCCG 178
Db	74 CCGAGCTGGGCACTCCGTGGCGAGCGCATCGCGGCATCGCGCACTGTGCGCGCGCAG 133
Oy	179 CGCACACCCGGGTGACAGCTGTGCGGCGCTGTGGTGTGGCGCGCGGCTTCATGACTGCACA 238
Db	134 CGGACGGGGCGCTGCATCGCGGGCGAGATGCTGCGCGCGGCTTCATGATACCAATG 193
Oy	239 CCCACGACGACAACTACTGCTCAGGCGCTCGCGCATGACGCCAACAGATCTTGCAGGGCG 298
Db	194 GCCACATGACCTGATTGTTCTGTAAGAAAACCGGGCGCTGGAATGAAAGCACACGAGGCA 253
Oy	299 TCACACAGGTGTCTACGGGCAATTCCGGCATATAGCTTGGCGCGCTGGCGACGCCAAC 358
Db	254 TCACGTGCGGTGTGTGCGCAACTCGGCATAGGGCGCGCGCGCGCGCGCTGCGCGGCA 313
Oy	359 CGCCCGCCCCCTGAGACCTGCTGAGAGAAAGGGGCTCTTACGTTTGAAGCGCTTCGCGG 418
Db	314 ATACGCGCG---GGCGTTGGCCCTGCTGCTGAGGCGCATCGCGCTGTTCTCGGACATGGCCA 370
Oy	419 ACTACCTGGAGCGGTGGCGGGCGACCGCGCGCGCTCAAGCCGCTGTATGTGGGGCG 478
Db	371 TGTATCTGGCGCGCTGGAGCGCGACGGCGCGCATGATCAAAGTGGCGCGCGCTGTGGGGCG 430
Oy	479 ATTCAACGTGTGGCGCGCGGCTCATGCGGATTTGCATGCGCGCGCGCACCGACGAGAAA 538
Db	431 ACGCCAACTCTGGCTGGCGCGCATGCGGATTCGGCGGCGACAGCCACCGCGGAAGGAC 490
Oy	539 TCGCGGCATCGGGAGACCTGCGCGAGAGAAGCATGGCGACGCGCGCATCGGCATTTGCA 598
Db	491 ACGGGCGATGAGGCGCATGCTGCGCGACAGCGCGCTGGAGGCGCGCGGTGGGATTCAGCA 550
Oy	599 CCGGCGCTTCTACCGCGCGCGCGCGCGCACCAACCGAAGATCATCGAGGTGTGCG 658

QY 92 ACGGAGCAACACCCCGGCGCGCCGACCTGGGCGCGGACGATCGCCG 151
 DB 10473 ACGGCTCCGCGGCGACCTCTACCGAGCGGACGTGTCACGCGCGCGATGCTT 10414
 QY 152 CCATGGGCGATCTGTTCGACGCCCGCGCACACCCG-----GGTCG 193
 DB 10413 CCATGCTCAAGAGAGGCGCGCGCGCGCTGCCAGCGCGCCGAGGACGCGGAGCTGG 10354
 QY 194 ACGTGTGCGCGCTGGTGTGGCGCGCGCGCTTCATGACTCGACACCCGACGACACT 253
 DB 10353 ACGCGAGGCGCTGTCTCTCTCCCGCTTATGACATGACAGCCACGACGACCTGG 10294
 QY 254 ACGTGTCAAGGCGCGCGACATGACGCCGACATCTCGAGGCGCGTACACGACGATGTC 313
 DB 10293 CGCTGTGCGGAGCGCGGACGACAGCGGCGCGCCGCGCGCTGAGAGTCA 10234
 QY 314 CGGGCAATTGCGGCGATCACTTCGCGCGCGCTGGCGCACGCCAAC----- 358
 DB 10233 TCGGCGAGAGCGCGGTGTGTACGACCGGTGACGACCGCTCGGCGAGAGTGGCC 10174
 QY 359 ---CGCGCGCGCGCGCGCGCTGTGAGCGAGAGCGCGCTTACCGTTTCGAGCGCTTG 415
 DB 10173 GGGCGATGCGCGGCGGAGAGCGCGCGCGCGCTTCAACCGCGCTGTATGTTGG 415
 QY 416 CGGACTACCTGAGCGCGTTCGCGGCGCACCGCGCGCGCTCAACCGCGCTGTATGTTGG 475
 DB 10113 GCGAGTACCTGAGC---CGGCTCGACCGCGGCGCATCGCGGTCAACCGCGGTACCTGATCC 10057
 QY 476 GCGATTCAACGCTGTGGCGCGCGCGGTTCATCCGAGCTTGCAGCGCGCGCGCACCGAGAG 535
 DB 10056 CGCAGAGCGAGGTGTGGCGCGCGCGCTCGCGCTGTGAGTGCAGCGCGAGGAGCGAGCGCG 9997
 QY 536 AATGCGCGCGCATGTGGCGCGCGCTGCGGAGAGAGCATGCGCGCGCGCTGCGCATTT 595
 DB 9996 ACGTGTGAGACATGCGCGCGCGCTGTGTCGCGGAGCGCTGAGAGAGCGCGGTGCGCTGT 9937
 QY 596 CGACCGCGCGCTTCTACCGCGCGCGCGCGCGCGCGCACCGACGAGAGATCATGAGTGT 655
 DB 9936 CGTGGGCGCTGACATACAGCGCGCGGATGTACGCCAGAGCGCGCGTGCAGCGAGCTGT 9877
 QY 656 GCGCGCGCGTGTGAGCG 715
 DB 9876 GCGGCGTGTGTGCGTGTACGCGCGCGCTACTGCGCGCGCGCGCGCGCGCGCGCG 9817
 QY 716 ACATGCG 775
 DB 9816 GCGCGCTGAGAGCGTGTACGAGAGATGTGTGAGCTACCGCGGAGCGCGCGCTGCGCTGC 9757
 QY 776 TGATCTCGACCAACAGTGTATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 835
 DB 9756 ACGTGTGCG 9697
 QY 836 CGCTATGTGAGCG 892
 DB 9696 CCCTCTGTGAGCG 9637
 QY 893 TGGCG 913
 DB 9636 CCG 9577
 QY 914 ACGAGACCGCGCGTGTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
 DB 9576 ACGAGATCTGTGAGAGCG 9517
 QY 974 AACTGAGCG 1003
 DB 9516 AGGTATGTGCG 9457
 QY 1004 ---CGCGCGAGCG 1033
 DB 9456 CCGGAGCG 9397
 QY 1034 ---TGCTGCG 1051

DB 9396 GCGTCCGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111
 QY 1052 CG 1111
 DB 9336 CG 9277
 QY 1112 GCG 1171
 DB 9276 GCGTCCG 9217
 QY 1172 GCG 1231
 DB 9216 ACGGCG 9157
 QY 1232 AGCG 1291
 DB 9156 AGGAGTGTGCG 9097
 QY 1292 GCG 1351
 DB 9096 GCTGTGTGCG 9037
 QY 1352 ATACCG 1411
 DB 9036 CG 8977
 QY 1412 ACG 1471
 DB 8976 ACGGCG 8917
 QY 1472 CAGCG 1494
 DB 8916 GCGGCG 8894

RESULT 12
 AF453501/c
 LOCUS 82746 bp DNA linear BCT 18-JUN-2002
 DEFINITION Actinosynnema pretiosum subsp. auranticum maytansinoid antitumor agent ansamitocin biosynthetic gene cluster I, partial sequence.
 ACCESSION AF453501
 VERSION AF453501.1 GI:21449342
 KEYWORDS
 SOURCE Actinosynnema pretiosum subsp. auranticum.
 ORGANISM Actinosynnema pretiosum subsp. auranticum; Actinobacteridae; Actinomycetales; Pseudonocardiales; Actinosynnemataceae.
 REFERENCE 1 (bases 1 to 82746)
 AUThORS Xu,T.W., Bai,L., Clade,D., Hoffmann,D., Toelzer,S., Trinh,K.Q., Xu,J., Moss,S.J., Leistner,E. and Floss,H.G.
 TITLE The biosynthetic gene cluster of the maytansinoid antitumor agent ansamitocin from Actinosynnema pretiosum
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (12), 7968-7973 (2002)
 MEDLINE 22056096
 PUBMED 12060743
 REFERENCE 2 (bases 1 to 82746)
 AUThORS Xu,T.W., Bai,L., Clade,D., Hoffmann,D., Toelzer,S., Trinh,K.Q., Xu,J., Moss,S.J., Leistner,E. and Floss,H.G.
 TITLE Direct Submission
 JOURNAL Submitted (28-NOV-2001) Department of Chemistry, University of Washington, Box 351700, Seattle, WA 98195-1700, USA
 FEATURES
 source
 1. 82746
 /organism="Actinosynnema pretiosum subsp. auranticum"
 /strain="ATCC 31565"
 /sub_species="auranticum"
 /db_xref="ATCC:31565"
 /db_xref="taxon:42198"
 <1. 82653
 /note="Cluster I of the maytansinoid antitumor agent ansamitocin biosynthetic gene cluster."


```
gene
CDS
    <1..1678
    /gene="asm37"
    <1..1678
    /gene="asm37"
    /codon_start=2
    /transl_table=11
    /product="unknown"
    /protein_id="AA054115.1"
    /db_xref="GI:21449383"
    /translation="IVYGLVYLAATGVLAAPASAAITSPVYSTATVTEVGRCT
    PARSRYVTDKRSATGYAOGGIGADHLNGLIRHSGGMATQVRYPPFSASGY
    ARMTVERADLGEQASPNVDVYQVEAPVETSAKVHTYTAGAGSVAYOPSSANPN
    PERGLYHHTGDCDADLSADLRRTYRTEOGITLAMCYFLAEHRSSPIAQTALDOR
    OFDARVAGVLYLRFAYTASESGEDAKERVLAHIDQLAPLLRANADYAAVYAGLR
    GAMGEWYTKNEGAGVSAANDANRKAAYDKLITLAPADRAVOLRTPKFKRALYGGG
    STRIGHNDCFLASDTPGTYEDPAPETPLAETRTPLMGECATVNPRTCCPTAF
    AELSRHMYTLNTDYOQDVLKANSAGCLEVYAKKLRITLDSVPTTTTTRAAIP
    VRNLTNDGYAAAANQRYNLVLRNTTGALTRLPLADPRTMLPGARALADQIIP
    SLATGOYALLLELLEDPALPGRPREYSIOTANTGMDARTGINDLKOTITVS"
    1825..2142
    /gene="asm36"
    1825..2142
    /gene="asm36"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AA054114.1"
    /db_xref="GI:21449382"
    /translation="WYTRSNVLCARAGCGAPACATARRPFCSTACKOAYRLRRR
    GSASHEDEQLRELLALVAKARTLTETAPPSAQATALVGLARRVEELAAADAGY
    SGG"
    complement(2129..3307)
    /gene="asm35"
    complement(2129..3307)
    /gene="asm35"
    /codon_start=1
    /transl_table=11
    /product="bicyclomycin resistant protein"
    /protein_id="AA054113.1"
    /db_xref="GI:21449381"
    /translation="MAVGKADSAVGGMPLALLIYALTGITIEIDMSVSPFDIARELG
    VGGPVOITVTLNPLFGICLALCHGRPLSDRGRPRVLLGSSAWMLYGLALPAGTCG
    LLTTRFVGGIGASAAVLYFTIIGDRIRGRPALMGLVYNAASALMTCAPILGGAL
    NRKAGMGNLYAVCTATSLILVFLPLPETRTARTATITARSVLADIRLLGSAVL
    NASTPLSLFAVLYEFAIASPELYTAAGLGCTGPAHLHLYVVASFAVPAFGAIIIP
    LLGPGAGVRCMGLAVVVGVLGFLTLDGCPAIVAGSVALFCVGFVAVCYPPVDFDSMSY
    PELAGAASSATMGRLALLVTLTLTALAGALSTGPFAPAAVAAVAVVLLSALSPR
    "
    complement(3450..4064)
    /gene="asm34"
    complement(3450..4064)
    /gene="asm34"
    /codon_start=1
    /transl_table=11
    /product="transcriptional regulator"
    /protein_id="AA054112.1"
    /db_xref="GI:21449380"
    /translation="MARRNPERRALLLDAAEVLAGEGARGLTFRAVDORAGVPAGTA
    SHYFAGRAETITGGERVYORLLDOAPLGEPTDGPDRSRVTELMHAIVDRYSAPTG
    FLALIEIRLEARNPELREVLTAARADIRFNAHEHNAASGLPGDGTVALMLALNML
    IIERLTLPDLTDHQRDLVTVLAVDRLLAGORPEKSPDESCK"
    4151..4723
    /gene="asm33"
    4151..4723
    /gene="asm33"
    /codon_start=1
    /transl_table=11
    /product="dihydrofolate reductase"
    /protein_id="AA054111.1"
    /db_xref="GI:21449379"
    /translation="MRKLYFFIAASIDGVIAEGEGRHDFPTAEPVWEHLRAEYET
    PTHVRPLGMLDTPNRRYDVLMGSGVOPALDAGVSGPYAHLROYVVSRLPADP
    AVALVRADPYGLVRELKREDGLDVMVLCGGADLAGQLPEIDELVYVTVVAVAGGTPPL
```

```
gene
CDS
    FRAGAPRAFRPVDYRAFDHGGVTTYRPA"
    complement(4854..5057)
    /gene="asm32"
    complement(4854..5057)
    /gene="asm32"
    /codon_start=1
    /transl_table=11
    /product="unknown"
    /protein_id="AA054110.1"
    /db_xref="GI:21449378"
    /translation="WITQONARIPARRSSGSRNARCACVAMPRQPSATRAQPRS
    SNAQDRASRRDRPRTSLNSA"
    5415..6461
    /gene="asm31"
    5415..6461
    /gene="asm31"
    /codon_start=1
    /transl_table=11
    /product="RNA polymerase sigma factor"
    /protein_id="AA054109.1"
    /db_xref="GI:21449377"
    /translation="WIGYSADGRRRRPVMGVTAEFTMETDEDIKIILATLRDAEA
    ANADSTRAYIKAGIRGRLTADQVEYLARVEAGLYAEHLIDLPAPARGSPVELR
    RYVRDQRAKHDLLENLPLVYSIAKRYCKGMALIDYQEGNLGVAVAEKEDYAKG
    FKSTYATWIRKQALITRALDOSRTIRLPLVHVEYVTKLARIERDLIRNREYATAE
    VAVELDVTEKVELKOHARTPSLIDQITGADVDSIGDLIDENAVAPLEVVSSALL
    SADLRVLELTLPREAGVYRMRYGPDVGLTLDDEVGVYTRERIRIOESRTAKL
    OHPARSGLEGYLD"
    complement(6501..9518)
    /gene="asm30"
    complement(6501..9518)
    /gene="asm30"
    /codon_start=1
    /transl_table=11
    /product="cytochrome P450"
    /protein_id="AA054108.1"
    /db_xref="GI:21449376"
    /translation="WVATGTRIPGKPLPLVGNLIDVLTSDLTVDVDFLDRCRHEHG
    IVALTFGQROVFASSSHELVAAMCSDPSGKGAHVPLAEVQKRFAGGGLTAGDEBNW
    GKARLMLPAFGPTAMRDHPAMDIDAEQMLVRRFRFGDHRIDVADQTLRLPLTIA
    LCAFGARFNFYDRAPHEVDARVSLVAGEAEERLPGVFFLGRNQRVYDIAETM
    NRIADGIVARALPAGERPDLEBMLCADPVTGERTSARVRYOLATFLIAGET
    TSCILSPAVRILAPHEVLPKAKDAVDGYLGRVPAPFEDLARLDYGOVLRFTLRHP
    TAPAFALAPPEPELGGHAIAGEPVLVNLPTLHRDPARWRDPDVPDRPAPERDE
    ITPCAMPPEFHGAKCIGREFALQEAIVLALVLFNFDLADLPDRRLTKOTLTLP
    DSLVARRAPVADLDAVATVETVPPQVATRHGTPPLHVFSGNSGSEGLARTIAG
    DGAARGWATVAPLDAVRAVLPASGVPVYSSYNGCAPDMAHEFRMTLOQPDLSG
    VDYVLGCGNLDMSATYORVPTLIDEMAAGARLREGATADRADPFGDWRMYEP
    LMPLLSACGCVGEGITGPRRVEVSDAAGIGDLASAVYLENEIYRGDAGSKRIIE
    LRLPDGTSYRTGYTSLTPQNHDPVRRAYARLGTAEVYVYESSAPGLVPGYRAL
    RVDELTRCYVLSAPAGAVYARLAEKPCPEERALATYGTGLTELELRPSCAVD
    LALALELPAVRTLYSISSAEORAEVALTVSVYSGYLSRVLPDRVAVGIAS
    PPESEFRPADNTVPVVLIAAGTGIAPFRGFLRAALAGGPPALLFCCRPELDDL
    YAEFFALGMLLEVDRAYSRRPGDEVHVOHRIAMORDRVRELVDAGARVYLCGDATR
    VGPVAEVIGRIGPGAGMDALRAGGYATDVE"
    9664..10245
    /gene="asm29"
    9664..10245
    /gene="asm29"
    /codon_start=1
    /transl_table=11
    /product="transcriptional regulator"
    /protein_id="AA054107.1"
    /db_xref="GI:21449375"
    /translation="MSAAVRREQLVAAALRYMKRDSITAAATYTAICAEDMPHGAHY
    CFRSKOELYTALTTGINTSIDAMPISDPLAVNVNRTMLAFAFTACQADPEALY
    LGELVTLAOREPELRELPAMOHREYVAKVAHAHLRLALEAGFGYSYDRALLAELVYVS
    LDGYASSMESHREGLAKASLDQFAVVFALVTR"
    10393..10857
    /gene="asm28"
    10393..10857
    /gene="asm28"
    /codon_start=1
```

```

/transl_table=1
/product="unknown"
/protein_id="AAM54106.1"
/db_xref="GI:21449374"
/translation="MTDTTRHTRPVONAAALAVGLVITAGFVPLTNTYDLEF
AGHSDALGVFVSAALNHLVLPAGLACTAGLSTAPRARARFLGGGLYAVLML
KGVYVDEHSTANPPLNADWMLHLGLAVGMITGLAALGRNAHTRARSDD"
complement(10912..11937)
/gene="asm27"
/complement(10912..11937)
/gene="asm27"
/codon_start=1
/transl_table=1
/product="glucose 1-dehydrogenase"
/protein_id="AAM54105.1"
/db_xref="GI:21449373"
/translation="MPRLAEQVVVVVVGASSGIGRTALRFAAGARVCAARNTAL
DGLVEVAGAGRAVAVTADIADFAVRAVAVRERGRDVTVMNAGICGVGRVD
TPAGEFDVMNVNLTGHVHGAALPALRRAGSVLIGVASVLELRSAFLDAPYAAAK
Query Match      6.6%; Score 116; DB 1; Length 82746;
Best Local Similarity 44.2%; Pred. No. 0.00011;
Matches 758; Conservative 0; Mismatches 935; Indels 21; Gaps 6;

QY 56 CCTTGACCTGCTGCTGCGGGGCGACCCCTCATGACGCGCAGCAACCCCGGGGGCGGC 115
DB 73869 CCGACCACTCTCTACACGCGGGGCTGCGCGCGCGCTGGGCGCTGCACGTCTCTCG 73810
QY 116 GCGCGACCTGGGCGTGCAGCGGCGACCGCATCGCCCGCATCGCGATCTGTGACGCGC 175
DB 73809 CCGCGCGGACACCGCGCTCGCGCGCGACGAGTGCACCGCGCTCGACGCGCTCGCG 73750
QY 176 CCGCGCACACCGGGGTGCACGTGTGCGGCGCTGTGTGTGCGCGCGCTTTCATCATCGC 235
DB 73749 CCGCGCACCGCGGGAGCGGGGACCAACCGCGGCGCTGCCCGCGCGCATGTCGCGG 73690
QY 236 ACACCCACGACGACAACTGCTCAGGCGTGCAGCATGACGCGCAAGATCTGCGAGG 295
DB 73689 ACCCGCGCTGCGCGCGCGCGCGTGTGCGGCGACGCGCGCGCGCGCGCGCGCGC 73630
QY 296 GCGTCACACGCTGTACAGGGCAATTTGGGCGATCAGCTGCGCGCGCTGCGCGCAACG 355
DB 73629 GCGGACGTCGCGCGAGGCGCG-----GCGCTGCGCGCGCGCGCGCGCGCGCGC 73576
QY 356 ACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
DB 73575 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73516
QY 416 CCGACTACCTGAGCGCGTGTGCGGGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
DB 73515 CCGAGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73456
QY 476 GCGATTCAACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGG 535
DB 73455 GCGACTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73396
QY 536 AATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 595
DB 73395 CCGCGAGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73336
QY 596 CGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 655
DB 73335 CGCTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73276
QY 656 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGC 715
DB 73275 CCGGACGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73216
QY 716 ACATGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 772
DB 73215 TCGCGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73156
QY 773 TGGTGTCTCGACCAAGATGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 832

```

```

DB 73155 CCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73096
QY 833 TGCGGCTATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 892
DB 73095 AACCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 73036
QY 893 TGGCGGCTTCACCATGCTCAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 952
DB 73035 GGGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72976
QY 953 CCTGTGCAAGCGCTTCCCGCACT--GACGCGCGCGCGCGCGCGCGCGCGCGCGCG 1009
DB 72975 CCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72916
QY 1010 AGCGCGCAATTCGAAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1069
DB 72915 GGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72856
QY 1070 TCATGATGACGAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1129
DB 72855 TGCGGCTGGCGTGTGCTGATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72796
QY 1130 CCGAGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1183
DB 72795 GGCACGTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72736
QY 1184 CCGCGGCTGTGGGCGCATATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1243
DB 72735 TGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72676
QY 1244 GGAAGATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1303
DB 72675 TCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72616
QY 1304 CCGGCTACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1363
DB 72615 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72556
QY 1364 TCGAACACCTTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1423
DB 72555 TGTGTGTC--GACAGCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72499
QY 1424 TCTGCGAAGACGAGCGTTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1483
DB 72498 CGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72439
QY 1484 CCGTGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1543
DB 72438 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72379
QY 1544 CAACCGTGAAGCGCAACCGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1603
DB 72378 TCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72319
QY 1604 ATATCTGGGCAAGAACTGATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1663
DB 72318 CCGGCACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72259
QY 1664 TGACCGTGAAGACCTGTGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1723
DB 72258 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72199
QY 1724 AGCTGGGCGAGATCGCGCTGAGTACGAGAAGCT 1757
DB 72198 CCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 72165

RESULT 13
SC6A11 28751 bp DNA linear BCT 12-MAY-2002
LOCUS Streptomyces coelicolor cosmid 6A11.
DEFINITION AL138978 AL645882
ACCESSION

```

```

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
TITLE
COMMENT
FEATURES
source
misc-feature
note

```

misc_feature
Nominal overlap with Streptomyces coelicolor cosmid 8F4
complement(6..872)
/note="Pfam match to entry PF01515 PTA_PTB, Phosphate
acetyl/putaryl transferase, score 460.50, E-value
1.4e-134"

RBS
RBS
gene
CDS
complement(1975..1978)
2103..2107
2118..2143
/gene="SC6A11.02"
/note="SC05426; pfkA2"
2118..2143
/gene="SC6A11.02"
/note="SC6A11.02, pfkA2, 6-phosphofructokinase, len: 341
aa; similar to many e.g. SW:059126 (PFP_AMYME)
6-phosphofructokinase from Amycolatopsis methanolica (352
aa) fasta scores: opt: 1274, z-score: 1463.6, E(): 0,
59.1% identity in 342 aa overlap. Also similar to
trkN:SCB51967 (EMBL:AL109661) Pfk, 6-phosphofructokinase
from Streptomyces coelicolor (342 aa) fasta scores: opt:
1251, z-score: 1437.4, E(): 0, 56.7% identity in 342 aa
overlap. Contains two Pfam matches to entry PF00365 Pfk,
phosphofructokinase and Prosite match to PS00433
phosphofructokinase signature"
/codon_start=1
/transl_table=11
/product="6-phosphofructokinase"
/protein_id="CAB72402.1"
/db_xref="GI:6941968"
/db_xref="SWISS-PROT:O9L1L8"
/translation="MRIGVLRAGDCCPLNMYVIRSVHRAVDNGDVIGPEDSVAGL
LDGRRLADLNAGSITLARGSTILLSSNLEDRIRREACENAGDIONFGDALPIGG
ESTLTPARRRLSDGIVLVGVPPTDNDISTDRKTFGEPTAVGATEAMDKRTTAESH
QRMYVAVNGHAGWIALESNGAAGAHGICLPERPFPADLVKNEERFSNGKFAVY
CAEAGVHPPEGSMDYKGAIDKFKHERFGIGTALAELERLRGRKEAPVILGVHVGSG
GVPTADRVLRTRFGMHAVEAHRGDFGRMTALRGTDVYVPLAEVTELTWPKDRM
DVAESYF"

misc_feature
2121..2804
/gene="SC6A11.02"
/note="Pfam match to entry PF00365 Pfk,
phosphofructokinase, score 278.20, E-value 1.6e-101"
2883..3017
/gene="SC6A11.02"
/note="Pfam match to entry PF00365 Pfk,
phosphofructokinase, score 52.20, E-value 9.9e-19"
2913..2969
/gene="SC6A11.02"
/note="PS00433-phosphofructokinase signature"
complement(3153..3866)
/gene="SC6A11.03c"
/note="SC05427"
complement(3153..3866)
/gene="SC6A11.03c"
/note="SC6A11.03c, unknown, len: 237 aa; no significant
database similarities"
/codon_start=1
/transl_table=11
/product="hypothetical protein SC6A11.03c"
/protein_id="CAB72403.1"
/db_xref="GI:6941969"
/db_xref="SPTREMBL:Q9L1L7"
/translation="MMRVRRPSPGSGSQSPENLAARRLRALNMGCEHYAHGIRVS
YGLPVPGLPVIAMERGTALEGPETLALAGVMSPELIGRPOSLEHRYAVAGLADL
EDVARRVGLPVIAMERESDDMRGSDRSALAVYLCTLPDFTVYGRREKLAADLIL
TSVTRRMGCGVRPIAKLAPLDRVLENTVQLHDQYGRMAATLSWGCAPDSSAGG
RDLPLDRLVDFHTRVRRHP"

gene
CDS
complement(3978..4868)
/gene="SC6A11.04c"
/note="SC05428"
complement(3978..4868)
/gene="SC6A11.04c"
/note="SC6A11.04c, possible integral membrane transport
protein, len: 296aa; similar to many proposed transport
proteins e.g. TR:Q9X9Y6 (EMBL:AL096743) putative integral

[illegible]

gene
CDS
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK25406.1"
/db_xref="GI:13425162"
/translation="MWLISAHKGECEPSEGVFNVLSSCMWGLYLAAPCAROPDPTRPD
WGRERFRRRSRSGTRRSRGPAPPAQAPAPAAAPAAQAPADQAPAPAAAPAEAPAPAD
PYVALVGPPTPTDEFTFLDLVDLQKCPKPVAGDEPAQLVSKSGFKKREOWFALEK
TEYAVLNGPNTNKNKVCVTITIDYQAPQAQELANGLMDAKTWENSPQLRLIRNDOTVG
SDPFRFTVMDADWAGGRAGLVYRKLKIDETSQKNFERAOLLYSTSR"
complement(1018. .1278)
/gene="CC3445"
complement(1018. .1278)
/gene="CC3445"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, epsilon subunit"
/protein_id="AAK25407.1"
/db_xref="GI:13425163"
/translation="MAKLHESLVAPEBELSEGEVDYQVARGAGDEGCVLANHAPFMTT
LREGVYTKDCAATKRVFDIOGCFADVGPEGLITLAHAVERA"
complement(1278. .1529)
/gene="CC3446"
complement(1278. .1529)
/gene="CC3446"
/note="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK25408.1"
/db_xref="GI:13425164"
/translation="MTEDADLANFCDELAENAVSPDANAGLIEHDSGLILNLGEEERE
BEGTEVLVATIRFCPCFCEGTETQEDIEALGVEETFN"
complement(1529. .3148)
/gene="CC3447"
complement(1529. .3148)
/gene="CC3447"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, beta subunit"
/protein_id="AAK25409.1"
/db_xref="GI:13425165"
/translation="MAKTPAEKPATPAKAPPAKPAKAAAPKAAAKAPAAKAPAKAKRR
APKAAAPKAPKPAAGLEGATGVLQIIIGAVVDIEFGALPAIINAVEVNTATOCRLVAV
PEVAHIOGQSYRAIAMDATGIVGQVPRTPGPIRVPGPGLTGRIMVIGPIEDIE
QGPISDSIRTIHRDAPSFABQNTAELVYTGIKYIDLMCYTTGGKGLGLGAGVCKR
IYTMQELINNIAKAGYGVSLVAGVERREGEDLHEHIESVANVDFRANNGSTEGSS
CALVVGOMNEPGRARARALTGLSIAEYFRDEGCDVLLFVNDLFRFOAGAEVSALL
GIIPAVAGOPTIATEMGNLQERISTKSGTSVQAIYVPADLTDPAPTSFADHL
ATTVLSRDIADAIIPAVDPDSTSRIMDPITIGEHYTVARVOEVLOQYKAKDIIID
AICDELSEEDKIIYVARARKIORTLSQPEHYABQFTNPGAPVQLKQTIINSFQYVD
GETDHPREGAEFTWGPPIEAVAKAKMAAEA"
complement(3196. .4071)
/gene="CC3448"
complement(3196. .4071)
/gene="CC3448"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="ATP synthase F1, gamma subunit"
/protein_id="AAK25410.1"
/db_xref="GI:13425166"
/translation="NASLKEMRNRISSVATOKITKAMQVYAAAKLRISODAAESAR
YARARLAVTANLAAVSGDGAPKRLAGTGRDRRHVVVAADRGLAGGFTSYIARAAR
AHIDGLIAQKQVRYVYCGKKTATOLARPYAGRIYETPDDLSYTRDLTISVAPRLADYI
TREYEGEDVDVTLFYSRFKSVVQOITPGLQIIPAVEGGEASGPTAVTEIEPSESE
IETLELPRMLTVOIILSALDDNNAAGFYASQMTAMDNATRNAGDMIKRYTLENSRQAA
IKRELLEITSGAEAV"
complement(4106. .5638)
/gene="CC3449"

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2003, 07:44:25 ; Search time 87 Seconds

(without alignments)
8893.123 Million cell updates/sec

Title: US-10-009-782-1

Perfect score: 1758

Sequence: 1 gaattccacttgatgcgcga.....ccctgagctagagaagctt 1758

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 389086 segs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US07_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	4.9	88421	9	US-09-976-059-1
2	81.8	4.7	1248	9	US-09-860-846-7
3	81.8	4.7	1248	10	US-09-861-289-7
4	81.8	4.7	13613	9	US-09-860-846-3
5	81.8	4.7	13613	10	US-09-861-289-3
6	73.8	4.2	15872	9	US-09-860-846-1
7	73.8	4.2	15872	10	US-09-861-289-1
8	73.4	4.2	4257	9	US-09-825-288A-1
9	72.4	4.1	3189	10	US-09-815-242-4056
10	72.2	4.1	13842	9	US-09-860-846-30
11	72.2	4.1	13842	10	US-09-861-289-30
12	72.2	4.1	36778	9	US-09-860-846-5
13	72.2	4.1	36778	10	US-09-861-289-5
14	69.6	4.0	11220	9	US-09-860-846-32
15	69.6	4.0	11220	10	US-09-861-289-32
16	67.8	3.9	1014	9	US-09-860-846-13
17	67.8	3.9	1014	10	US-09-861-289-13
18	67.8	3.9	13029	10	US-09-815-242-4052
19	64.2	3.7	2712	10	US-09-748-033-4

20	63.2	3.6	1764	10	US-09-815-242-7984	Sequence 7984, Ap
21	62.2	3.5	804	9	US-09-773-748-2	Sequence 2, Appli
22	62.2	3.5	1290	10	US-09-815-242-7834	Sequence 7834, Ap
23	61.6	3.5	88421	9	US-09-976-059-1	Sequence 1, Appli
24	60.8	3.5	1185	10	US-09-887-576-784	Sequence 784, App
25	60.8	3.5	2787	9	US-09-860-846-40	Sequence 40, Appli
26	60.8	3.5	2787	10	US-09-861-289-40	Sequence 40, Appli
27	60.4	3.4	4689	9	US-09-860-846-34	Sequence 34, Appli
28	60.4	3.4	4689	10	US-09-861-289-34	Sequence 34, Appli
29	60	3.4	7419	10	US-09-815-242-4009	Sequence 4009, Ap
30	59.8	3.4	1815	10	US-09-758-269-13	Sequence 13, Appli
31	59.8	3.4	1839	10	US-09-815-242-4006	Sequence 4006, Ap
32	58.6	3.3	1107	10	US-09-748-033-6	Sequence 6, Appli
33	58.4	3.3	1038	10	US-09-815-242-4048	Sequence 4048, Ap
34	58.2	3.3	2280	10	US-09-815-242-4005	Sequence 4005, Ap
35	57.2	3.3	1325	10	US-09-770-517C-12	Sequence 12, Appli
36	57.2	3.3	1677	10	US-09-770-517C-1	Sequence 1, Appli
37	57.2	3.3	1897	10	US-09-770-517C-25	Sequence 25, Appli
38	57.2	3.3	6854	9	US-09-922-683-7	Sequence 7, Appli
39	56.6	3.2	1533	9	US-09-712-363-81	Sequence 81, Appli
40	55.8	3.2	1374	10	US-09-815-242-7850	Sequence 7850, Ap
41	55.8	3.2	4826	10	US-09-772-304A-1	Sequence 1, Appli
42	55.6	3.2	1203	9	US-09-879-312-1	Sequence 1, Appli
43	55.4	3.2	1224	10	US-09-815-242-7867	Sequence 7867, Ap
44	55.4	3.2	1428	10	US-09-815-242-4000	Sequence 4000, Ap
45	55.2	3.1	798	9	US-09-813-453A-24	Sequence 24, Appli

ALIGNMENTS

RESULT 1
US-09-976-059-1
Sequence 1, Application US/09976059
Patent No. US20020164747A1
GENERAL INFORMATION:
APPLICANT: Farnet, Chris
APPLICANT: Zazopoulos, Emmanuel
APPLICANT: Staffa, Alfredo
TITLE OF INVENTION: Genes and Proteins for Biosynthesis of Ramoplanin
FILE REFERENCE: 3019-PCT
CURRENT APPLICATION NUMBER: US/09/976,059
CURRENT FILING DATE: 2001-10-15
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 88421
TYPE: DNA
ORGANISM: Actinoplanes sp.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2077)..(3078)
OTHER INFORMATION: ORF 1: positive strandedness
NAME/KEY: misc_feature
LOCATION: (3118)..(4032)
OTHER INFORMATION: ORF 2: positive strandedness
NAME/KEY: misc_feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3: positive strandedness
NAME/KEY: misc_feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3: positive strandedness
NAME/KEY: misc_feature
LOCATION: (4038)..(5048)
OTHER INFORMATION: ORF 3: positive strandedness
NAME/KEY: misc_feature
LOCATION: (6665)..(5814)
OTHER INFORMATION: ORF 4: negative strandedness
NAME/KEY: misc_feature
LOCATION: (7703)..(6693)
OTHER INFORMATION: ORF 5: negative strandedness
NAME/KEY: misc_feature
LOCATION: (9464)..(8130)
OTHER INFORMATION: ORF 6: negative strandedness
NAME/KEY: misc_feature
LOCATION: (9691)..(10761)

OTHER INFORMATION: ORF 7; positive strandedness
NAME/KEY: misc_feature
LOCATION: (12751)..(10829)
OTHER INFORMATION: ORF 8; negative strandedness
NAME/KEY: misc_feature
LOCATION: (13617)..(12802)
OTHER INFORMATION: ORF 9; negative strandedness
NAME/KEY: misc_feature
LOCATION: (15203)..(13614)
OTHER INFORMATION: ORF 10; negative strandedness
NAME/KEY: misc_feature
LOCATION: (15591)..(15863)
OTHER INFORMATION: ORF 11; positive strandedness
NAME/KEY: misc_feature
LOCATION: (15860)..(19035)
OTHER INFORMATION: ORF 12; positive strandedness
NAME/KEY: misc_feature
LOCATION: (19032)..(39713)
OTHER INFORMATION: ORF 13; positive strandedness
NAME/KEY: misc_feature
LOCATION: (39713)..(65800)
OTHER INFORMATION: ORF 14; positive strandedness
NAME/KEY: misc_feature
LOCATION: (65826)..(66530)
OTHER INFORMATION: ORF 15; positive strandedness
NAME/KEY: misc_feature
LOCATION: (66346)..(67370)
OTHER INFORMATION: ORF 16; positive strandedness
NAME/KEY: misc_feature
LOCATION: (67384)..(70059)
OTHER INFORMATION: ORF 17; positive strandedness
NAME/KEY: misc_feature
LOCATION: (70099)..(70662)
OTHER INFORMATION: ORF 18; positive strandedness
NAME/KEY: misc_feature
LOCATION: (70659)..(71906)
OTHER INFORMATION: ORF 19; positive strandedness
NAME/KEY: misc_feature
LOCATION: (73439)..(71964)
OTHER INFORMATION: ORF 20; negative strandedness
NAME/KEY: misc_feature
LOCATION: (74216)..(73563)
OTHER INFORMATION: ORF 21; negative strandedness
NAME/KEY: misc_feature
LOCATION: (73424)..(74213)
OTHER INFORMATION: ORF 22; negative strandedness
NAME/KEY: misc_feature
LOCATION: (75535)..(76464)
OTHER INFORMATION: ORF 23; positive strandedness
NAME/KEY: misc_feature
LOCATION: (78110)..(76449)
OTHER INFORMATION: ORF 24; negative strandedness
NAME/KEY: misc_feature
LOCATION: (79864)..(78107)
OTHER INFORMATION: ORF 25; negative strandedness
NAME/KEY: misc_feature
LOCATION: (81624)..(79861)
OTHER INFORMATION: ORF 26; negative strandedness
NAME/KEY: misc_feature
LOCATION: (81909)..(81682)
OTHER INFORMATION: ORF 27; negative strandedness
NAME/KEY: misc_feature
LOCATION: (82346)..(82062)
OTHER INFORMATION: ORF 28; negative strandedness
NAME/KEY: misc_feature
LOCATION: (82587)..(84446)
OTHER INFORMATION: ORF 29; positive strandedness
NAME/KEY: misc_feature
LOCATION: (84461)..(85548)
OTHER INFORMATION: ORF 30; positive strandedness
NAME/KEY: misc_feature
LOCATION: (85556)..(86845)
OTHER INFORMATION: ORF 31; positive strandedness

NAME/KEY: misc_feature
LOCATION: (87372)..(86803)
OTHER INFORMATION: ORF 32; positive strandedness
NAME/KEY: misc_feature
LOCATION: (87494)..(88420)
OTHER INFORMATION: ORF 33; positive strandedness; N-terminus only
US-09-976-059-1

Query Match 4.9%; Score 86; DB 9; Length 88421;

Best Local Similarity 43.1%; Pred. No. 7,5e-08; Matches 605; Conservative 0; Mismatches 775; Indels 24; Gaps 3;

35 TGTCCATTCGATTCCTCCAGCCCTTGCACCTGCTCTGCGGGGCGCACCTCATGACG 94
Db 46230 TGGCCCGGACAGAGTGTGCGGCACAGGTGTTCACGCTGCGGACGCGAGCGGTGCAAC 46289
95 GCAGCAACACCCCGGGGGGGGGCGGACCTGGGCTCCGGCGCCACCGCATCGCGGCA 154
Db 46290 ACATCTTGACCCCGCGGGGGGGCGAGCCCGTTCGGGTGAGCCTGCCCGCGGCGC 46349
155 TCGGCAATCTGTGAGACCGCGCGGCACACCCGGGTGAGCTGTGCGGCTGTGTCG 214
Db 46350 GGGTGGAGAGGGCGGTCCCGCAAGGGCGCGGTACGCTTGACCTGGCCCGGAGATCC 46409
215 CGCCCGGCTTATGATGATCGACACCAACCAAGACACACTTACTGCTAG-----GCGTC 268
Db 46410 CGCTCGTGGCGCTGCTGCTGCGCCCGCGGACGACACCACTGCTGCTGCTGTC 46469
269 GCGACATAGCGCCCAAGATCTGCGAGGGCGCTACCAAGGTGTGTCAGGGGCAATTCGCGCA 328
Db 46470 ACCACATGCGGCGCGACGAGGTGTGATGAGGGCGCTGCGCGGACCTGCGGACCGCTT 46529
329 TCAGCCTGGCGCGCTGCGACGCGCAACCGCGCGCGCGCGCGCTGCTGAGAGAG 388
Db 46530 ACGCGCGCGGGGGGGGGGGGCGAGCGCGCGCGCGAGTCCGAGACCTGCGCTCAGTACGCCG 46589
389 GCGGCTTTACCGCTTTCAGAGCGCTTTCGCGGACCTGAGACGCGCTTGGCGGCGACCGCGG 448
Db 46590 ACTACGCGCTTGGCAGCGGTGACCTGCTGCGGCTCGACAGGACCGCGGAGCTGATCT 46649
449 CGCGCGCTAAGCGCGCGCTGATGAGGGGCAATTCAACCTGCGCGCGCGGCTGATCGCGG 508
Db 46650 CCGGCGAGATGCGCCACGAGTGGCGGAGCGGCTCGACGAGGCGCGGAGAGTGAACCTGC 46709
509 ACTTCAGCGCGCGCGCGCACGAGAGAAATCGCGGCGCATGCGGAGCTGCGCGGAGAG 568
Db 46710 CCGCGGACCGG 46769
569 CCATGCGCAGCGCGCGCATTCGCAATTCGACCGGCGCTTACCGCGCGCGCGCGCGCG 628
Db 46770 TCCCGCGCGAGTGCACCGAGACCTGCGCGGCTGCGCGGACCGAGACCGCTCT 46829
629 CCACACCGAGAGATCATTCAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCATTCG 688
Db 46830 TCATACCTCTGAGGCGCGCGCGGCGGTCTGCTGCGCGCGCGCGCGCGCGCGCGCG 46889
689 CCACCCACATGCGCGAGGAGGCGAGACATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCA 748
Db 46890 TCCGATGCGACCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 46949
749 TCGCGCGCGAGCTGAGCTGCGCGGTGTGATCTGCGACACCAAGTCTGAGGCGCGCGCA 808
Db 46950 TCTTGTCAACAGCTGCTGTGCGCACCGACTGACCGGGGAGCGCGGCTGACCGAGCG 47009
809 ATTTCGCGCGCGCGCGCGAGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
Db 47010 TGTCTGCGAGGTGCGAGCTGACCTGCGGGCGCTGCGCGCGCGCGCGCGCGCGCTG 47069
869 TCTCGCTGAGCGCGATTCCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 928
Db 47070 AGAAGCTGTGAGAGAGCTCAACCGCGCGCGCTC-----GCTGCGCGCGCGCGCGCT 47123
929 TGTGCGCGGAGCGACCATCATCATCTGCTGTGCAAGCCCTTCCCGGACCTGAGCGGGCGG 988

Db	47124	TCGACGTGATGGTTCACCTCGAGACGGCGCGGCGCGGCGCGGACCGGACGGCGCCCGACGCTGCGCGGCGCTTGG	47183
OY	989	ACCTGAGTGAAGTCCGGGCGCGAGCGCGGCGAATTCAGATGACAGCTGGTGGCCCGAGCTGC	1048
Db	47184	CGATGTCCGTCGTGCGGACCGGCGCGCGCTGTTCCGGCCAGTGTGCACTCGACCTCAGCTCAGATTCA	47243
OY	1049	AGCGGCGCGGCGCGCCACTTACTCTTACATGATGATGAGCAAGAACCCGACGCTGCAGCGCATCTCGGCGT	1108
Db	47244	CCGAAACCTTTCGACCCCGCGGGGAGACCGCGCGCGCTCGCGCTGCACCTCAGCTCAGCGCGGCG	47303
OY	1109	TGCGGCCGACCATGATTCGGCTCGAGGGCGTTCGGCGCAGCAGACGAGCGCCCGCATTCGCGCC	1168
Db	47304	CCGACCTCTTGCAGCCGGGACAGCGCGCGCGCGCTGCGCGGCTACCTGAGACCGCGCTTCTCG	47363
OY	1169	TGTGGGGACCTTCCCGGGGGTGTCTGGGGGACATATGCGCGGACCTGGAGGCGCTTCCGCG	1228
Db	47364	GCGTCTCGCGGCGCATCCGGGGCGCGGCGCGCTTGGCGGAGGTGCA-----CCGCG	47411
OY	1229	TGGAACGCGCGGTATGGAAGATGACCGGCGCTGACCGCGCGCGCTTTCGGGCTTCGCGGCGG	1288
Db	47412	TGGAGGCGAGGAGAGACCGGCTCATGCTTCGCGCGGTGAGGAGCGCCGCGCGCGCGCTTGC	47471
OY	1289	GCGGCGACCTGAGGACCGGGTACTTCGCGCGACCTGTGTGTTCACACCGCGCGACAGGTGG	1348
Db	47472	CCGAGATACCGCTCGCGGCGCTCTGTCGCGAGACATGTGGCCCGCGACGCGCGGGTGGGCTGG	47531
OY	1349	CCGATACCGGCACCTTTCGAACACCTTCAGACGAGCGCGCGCGGCAATTCATTCCTGTACG	1408
Db	47532	CGGTACCGGAGACCGGACCGAGCGAGCTGACCTTACGCGGAGCTGACGAGGGGGCGCGCCGA	47591
OY	1409	TCGAACGGGCGCGGCTTCGCGACG	1432
Db	47592	TGCGCGCTGCGTGGCGCGCGCACG	47615

[illegible]

Db	335	TCACCCGGGTTCTTGCGGACATGAGACCCGGACACACGGGACAACTCGAACCCGGACACAGTGG	394
QY	212	TCGGGCCCCGGTTTCATTCAGACTTCGCACACCCACGACGACAACTACCTGCTCAGGCGTGGC	271
Db	395	CCGGCGGGGTATACACCCCGGACCTTCGGGCCCCCTCTCCGCTCCACCTGCTGGGGCCGCCCT	454
QY	272	ACATGACGCCCCAAGATGATCCGAGGGGCTGCACCGGTGGTCACGGGCAATTGCGGCATCA	331
Db	455	GGCGGCCCGGACACAGCTCTCGGAAGGTGCCGACGACGACGCTTGGCGCTGTACTTTCGACG	514
QY	332	GCCTGGCGCGCTTGGCGCACGCCAACCCGCCCCCCTCCGACCTGCTGTGACGAAAGCG	391
Db	515	CCGGGCAAGGCTCTGGCTGGCGGTGCAGCGGCGGCGCCCGCGGACAGCTCGCGGACGCGCG	574
QY	392	GCTCTTACCGTTTGGAGACGCTTTGGCGGACTACTGTGACCCGCTTGGGGGACACCGCGGG	451
Db	575	AGGCTTTTAGGTTTCA---GCGCACCAAGGCGGTCAACGCTCTTGACGAGGGCGGGCGCTGC	631
QY	452	CCGTACAGCGCGCTGTATGTGTGGGCGCATTCACAGCTGGCGGCGCGCGGATGATCCGAGCT	511
Db	632	TCACCGAGAGCGCGGACCTCTGCGCGCCGGATCCGGGCGCTTCACAACTTTCGGCTTCGAC	691
QY	512	TGCAAGCGCGCGCCACCGGACGAGAGAAATGCGCGCCATGCGGACCTGGCCGAGAGACCA	571
Db	692	TGCCCCGGGGATGCGCCCGCGCGGGAGACCAACGCCAAGATGAGCAGAGGCGCGCGCGGCCA	751
QY	572	TGGCCAGCGGGCGCATGCGGATTTTCACAGCGGCGCGCTTCATCAGCCGGCGCGGCGCGCGGCA	631
Db	752	TGGGCTCTACCTCCCTCGAGCGGCTTTCGGAAGTCAATCAGACCGGAACCGGCGGCAACACG	811
QY	632	CCACCGGAAGAGATCATATGAGAGTGTGCCGGCGCTGAGCGCGCATGCGGCGATTCAGGCCA	691
Db	812	CCGGCTTACCGCGAGACACTTGGCGGACCTCCCGGCGCTCTCTCCGACACGACGACCGC	871
QY	692	CCACATATGCGCGAGAGGCGAGACATCATGTGGCGCGCGCTGAGAGAAACCTTCCGATCG	751
Db	872	ACGGGCTCTACAAACACACCACTACGTATGTGTGAGATGCAGCAGGCGCACACCGGCAATCC	931
QY	752	GCGCGGAGCTGGAGCTGCCGGTGTGATCTCGACACCAAGGTCAATGGGCCAGGCCAATT	811
Db	932	ACCCGCGACTCTGATGAGAGTCTCTGAAGGCGGAAGGCTGCACACCGCGCGCTACTTCT	991
QY	812	TCGGGCGCTGGCGGAGAGCGCTGCCGCTGATGAG-----GCGGCATATGGCGCGC	862
Db	992	CGCGGGGCTGCCAGAGCTGAGAGCCCTTACCGGGGAGCGCGGACCGCCGCTGGCGACCA	1051
QY	863	AGGACGTCTGTGGCTGAGAGCGCTATCCCTAGGTGGCGGGCTCCACCAATGTGTCAAAGGAGACC	922
Db	1052	CCGAAACGCTTGGCGCGCGCGGTCTCTCTCCCTCGACACCGGACCGCATTCGGCGACGACG	1111
QY	923	GCGTGTGCTTGGCGCGGACGACCATCATCACTGTGTGCAAGCCCTTCCCGAATGAGCG	982
Db	1112	ACATCTCCCGGGTGGCGGACCTCTGCTGCTCTGTGGCGGACCGCGGCGCGCAACTGACCG	1171
QY	983	GGCGGACCTGAGTGAAGTCCGGCGGCGGCGG	1015
Db	1172	CGCGCCACCGCGACAGGCCCCCGCCCGCTCG	1204

RESULT 3
US-09-861-289-7
Sequence 7, Application US/09861289
Patent No. US20020110897A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.438051
CURRENT APPLICATION NUMBER: US/09/861,289
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537

Db	12132	TGGGGCTCACTCCCTCCGACGCGTTTCCCGAAGTATCATGACCGGAAACCGGGCGAACACAG	12073
Qy	632	CCACCGAAGAGATATCATCGAGGTGTCCGGGCGCTAGCGCCGATGGGGCATCTACGCCA	691
Db	12072	CGCGCTTACCGGGAGCAGCTCGGGGACCTCCCGGGCTCTCTGTCCCGACACGACCGCC	12013
Qy	692	CCACATCGCGGCGAAGAGCGAGCATCTGTGGCGGCGCTGGAGGAACCTTCCGATCG	751
Db	12012	ACGGCTTACACACCAACCAACGATGCTGATGTGTGAGATGACGAGGCGACACACCGGCGATCC	11953
Qy	752	GCCGCGAGCTGGAGCTGCGCGGGTGGTATCTGCACCAAGGTCTATGGGCGCCAGCCCAATT	811
Db	11952	ACCGCGACCTGCTATGGAGGTCTCTGAAAGGCCGAAGGGCTGCACACCGCGCTTCTT	11893
Qy	812	TCGGCGCTCGCGCGAGACGCTGCGCGTGATCGAG-----GCCGCGATGGCGCGCGC	862
Db	11892	CGCGCGGCTGCACGAGCTGGAGCGGTAACCGGGGCGACGCGGACGCCCGCGCGCGCACA	11833
Qy	863	AGGAGCTCTCGCTGGACGCGGTATCCCTACGTGTGCGCGGCTCCACCATGCTCAAGCAGAGAC	922
Db	11832	CCGAACGCGCTCGCGCGCGCGCGGTGCTCTCCCTGCCACCGCGCCGCTCATGTGGCGAGCG	11773
Qy	923	GCGTCTGCTGGCGCGAGCACACATCATACCTGTGTGCAAGCCCTTCCCGCAACTGACCG	982
Db	11772	ACATCGCGCGGGTCCCGCACCTGCTGCGCTCTTGTGCGCACCGCGCGCGCGCAACTGACCG	11713
Qy	983	GCGCGACCTGTAGATTCGCGCGCGAGCGCG 1015	
Db	11712	CGCGCACCGCGACACGCGCCCCCGCGCGCGCTCG 11680	

: RESULT 5
 : US-09-861-289-3/C
 : Sequence 3, Application US/09861289
 : Patent No. US20020110897A1
 : GENERAL INFORMATION:
 : APPLICANT: Sherman, D. H.
 : APPLICANT: Liu, H.
 : APPLICANT: Xue, Y.
 : APPLICANT: Zhao, L.
 : TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 : FILE REFERENCE: 600.438051
 : CURRENT APPLICATION NUMBER: US/09/861,289
 : CURRENT FILING DATE: 2001-05-18
 : PRIOR APPLICATION NUMBER: 09/105,537
 : PRIOR FILING DATE: 1998-06-26
 : NUMBER OF SEQ ID NOS: 43
 : SOFTWARE: FastSeq for Windows Version 3.0
 : SEQ ID NO 3
 : LENGTH: 13613
 : TYPE: DNA
 : ORGANISM: Streptomyces venezuelae
 : US-09-861-289-3

Query Match	Similarity	4.7%	Score	81.8	DB 10	length	13613
Best Local	Similarity	44.2%	Pred.	No. 4.3e-07			
Matches	439	Conservative	0	Mismatches	542	Indels	12
						Gaps	2
QY	32	CCATGTCCCCAATCCGATTCGCCAGCCCTTGACTGCTGCTCGGGGGGCGACCCATCATG	91				
Db	12669	CCACCTGCACAGCCCAAGGCGGGGCTTCACACTCTCCGCAACGGCGCGCTCACCGGG	12610				
QY	92	ACGGCAGCAACACCCCGGGGGCGCGCGCGACCTGGGCGTGGCGGGCGAACCGCATCGCG	151				
Db	12609	AAGTGTATCATGCGGTGATGAGACTTGTCGCCGACCCCGCAGCAGCATGCGCTGATCGGCG	12550				
QY	152	CCATGGGCGCATCTGTGCGAGCGCGCGCCGACACCCCGGGGTGACGCTGTGGGGCGCTGGT	211				
Db	12549	TCACCCCGGCTTTCGGCGACATCGAACCGGGACACCGCGCAACTCGAACCCGGACACAGTGG	12490				
QY	212	TGGCGCGCGGCTTCATCGACTGCGACACCCAGACGACCAACTGACTGCTCAGGCGTGGC	271				

Db	12489	CCGCGCGGCTACACACCCCGACCTCGCGCGCTGTTGAGGCTCACACTCTGTGGGCGCGCCCT	12430
Qy	272	ACATGACGCCCAAGATCTGGCAGAGGCGCTTACACAGGTGTGTACAGGGCAATTGGCGCATCA	331
Db	12429	GCGCGCGCGACACAGCTGGCGGAAAGTGTGGCGAGACAGCGCTTCGCGCTTACTTGTGAGG	12370
Qy	332	GCGTGGCGGCGCTGGAGGACAGCGCAACCGCGCCCGCCCTCGGACTCTGTGGAGAGGGG	391
Db	12369	CCGCGACAGCGCTTGGGCTGTGGGTGTGAGAGCGCGCGCCCGCGGAGCTTGGGCGACGGC	12310
Qy	392	GCTCTTACGTTTTCGAGCGCTTTCGCCGACTTACCTGGACGCGTTGGCGGCGACGCGGGCG	451
Db	12309	AGGTCCTTCACTTTCACA---CGCACCAAGCGGCTTCAAGCGCTTTCGAGAGGCGGCGCTGG	12253
Qy	452	CCGTCACAGGCGCGCTTATGTGTGGGCGCATTCAAAGCTTCGCGCGCGCGGCTATCCGGACT	511
Db	12252	TCACCGACAGCGCGCACTTGGCGCGCGGATCCGCGCTTCACAACTTGTGCTTGGAC	12193
Qy	512	TGCAGCGCGCGCCACCGACGACGAGAAATGCGGCCATCGCGGACTGGCGGAGAACCA	571
Db	12192	TGCGCGGCGGACCGCGCGCGCGGGGAGCCAAACGCCAAGATGAGGACGCGCGCGGCCA	12133
Qy	572	TGGCGACAGGCGCGCATTCGGCATTTGACACGGGCGCTTTCATCCGCGCGCGCGCGGCA	631
Db	12132	TGGGCGCTACACTCTCCCTGACGGGTTCCTCCGAGGTATGACACCGGACCGGCGCAACACG	12073
Qy	632	CCACCGAAGAGATCATCGAGAGTGTGCGGCGCGCTGAGGCGCATGGCGGATTCAGGCCA	691
Db	12072	CCGCGTACCGCGGACACATTCGGCGGACCTCCCGGGGCTCTGTGGCCGACACGACCGCC	12013
Qy	692	CCCACATGCGCGACGAGAGGCGACACATGTGGCGCGCTGGAGAGAACTTCCGCATCG	751
Db	12012	ACGGGCTTCACCAACACACACAGTACGTATGTGTGATGATGAGACGAGCGCACCGGATCC	11953
Qy	752	GCGCGGAGCTGGAGAGTCCGCGGTGTGATTCGCCACCAAGGTATGGGCGACGCCAATT	811
Db	11952	ACGCGGACCTGTGTATGAGAGGTCTCTGAAGCGCGAAGGGGTGCACACCCGCGCTTCTCT	11893
Qy	812	TGCGGCGCCTCGCGCGAGAGCGTCCGCTGATGAG-----GCCGCAATGAGCGCGC	862
Db	11892	CGCGGGGCTGCCACAGAGCTGAGAGCGTACCGGGGGCACCGCAGCGCCCTGCGCAC	11833
Qy	863	AGGACGTCCTGCTGAGAGCGGTATCCCTACGTGGCGGCTCCACCAATGCTCAAGCAGAGCC	922
Db	11832	CCGAAACGCTTCGCGCGCGCGGCGGTCTCCCTCGACCGAGCGCATGCGGAGAGGAC	11773
Qy	923	GCGTGTCTGGCGCGGACGACCATCATCATCTGTGTGCAACGCTTCCCGGAATGAGG	982
Db	11772	ACATTCGCGCGGGGTGGCGAGCTGTCTCGCTTTCGCGGACCCGCGGCGGAACTGACCG	11713
Qy	983	GCGCGACCTGATGAAGTGCGGCGGACGCGG	1015
Db	11712	CGGCGCACCGGACAGGCGCCCGCGCCGCTGG	11680

```

RESULT 6
US-09-860-846-1
; Sequence 1, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438US1
; CURRENT APPLICATION NUMBER: US/09/860,846
; PRIORITY FILING DATE: 2001-05-18
; PRIORITY APPLICATION NUMBER: 09/105,537
; PRIORITY FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1

```

LENGTH: 15872
 TYPE: DNA
 ORGANISM: Streptomyces venezuelae
 US-09-861-289-1

Query Match 4.2%; Score 73.8; DB 9; Length 15872;
 Best Local Similarity 43.2%; Pred. No. 1.3e-05;
 Matches 586; Conservative 0; Mismatches 747; Indels 24; Gaps 4;

175 GCCGCGCACACCGGGGTGACAGTGTGGGGCTGTGGTGGCGCCGCTTACATGACTCG 234
 4799 GGGGAGGAACTGTGGCGGGCGGACCCCGTGTGGCGCCCGCTGACACCGTGTACG 4858
 235 CACACCCAGCAGCAACTGTCTGTGAGCGGTGAGAGAGAGAGAGAGAGAGAGAGAG 294
 4859 GCCCTGACCGGTACCTGTGACCGCGCGCTGCGGAGATCTCCCGCGGGAGAGAGCTG 4918
 295 GCGCTACCAAGGTGTGACAGGGCAATTGGCGCATGAGCTGGCGCGCTGGCGGACGCC 354
 4919 GACCTACCGGTGTACACCGACCGCGCCCTTGTGCGCTTGTGAGGAGGCGGTGTCGCG 4978
 355 AACCCG 414
 4979 CTGGAACACCAAGCGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5038
 415 GCGGACTACCTGTGACAGCGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 5039 GCGCGGACGCTGCGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5098
 475 GCGCATTCACAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
 5099 GCGCGGCTCATGTGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5158
 535 GAATTCG 594
 5159 GCGGAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5218
 595 TCGACCGCGCGCTTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654
 5219 GGACCGACCGCGGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 5278
 655 TCGCGCGCGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 714
 5279 TGGCGGAGACCG 5338
 715 CACATTCG 774
 5339 CACATGTGAGCGAGTGTGTGAGTGTGTCTCTCGGGTCCCGGAGGGCTGTGAGTGT 5398
 775 GTGATCTGCAACCAAGTGTATGGCCAGCCCAATTGGCGCGCGCGCGCGCGCGCG 834
 5399 CCGCGGATCCCGGTGTCTCTCAACGGTCAACGGCGCGGTGTCTCTCTCTCTCTCT 5451
 835 CCGGTGTGAGCG 894
 5452 --GCTCACTGTGCGCGGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5509
 895 GCGCGGTGCAACGATGTGTCAACGAGCGCGCGGTGTGTGCGCGCGCGCGCGCGCG 954
 5510 GCGGTGCGCGACCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5569
 955 TGTGTGCAAGCGCTTCTCTCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1014
 5570 GTCTGTCAAGCGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5629
 1015 GCGCAATTCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1074
 5630 GACGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 5689
 1075 ATGACGAAACCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1134
 5690 GGTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 5749

1135 GCGCTGCG 1194
 5750 GGGCG 5809
 1195 GGGCAC-----TATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248
 5810 GAGGCGCGGTACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5869
 1249 ATGACCG 1302
 5870 ACCAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5929
 1303 GCGGCGTACT---TCG 1359
 5930 GACGACCG 5989
 1360 ACCTTGACACCGCTTACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1419
 5990 TTTCTGTGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6049
 1420 CCGGTGTGCAAGAGCGGTGTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1479
 6050 ACCCTGTGAGCG 6109
 1480 GCGCGTGTGAGCG 1516
 6110 AGCGGAGCGCGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 6146

RESULT 7
 US-09-861-289-1
 ; Sequence 1, Application US/09861289
 ; Patent No. US20020110897A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sherman, D.H.
 ; APPLICANT: Liu, H.
 ; APPLICANT: Xue, Y.
 ; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
 ; FILE REFERENCE: 600.438051
 ; CURRENT APPLICATION NUMBER: US/09/861,289
 ; CURRENT FILING DATE: 2001-05-18
 ; PRIOR APPLICATION NUMBER: 09/105,537
 ; PRIOR FILING DATE: 1998-06-26
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 15872
 ; TYPE: DNA
 ; ORGANISM: Streptomyces venezuelae
 US-09-861-289-1

Query Match 4.2%; Score 73.8; DB 10; Length 15872;
 Best Local Similarity 43.2%; Pred. No. 1.3e-05;
 Matches 586; Conservative 0; Mismatches 747; Indels 24; Gaps 4;

175 GCCGCGCACACCGGGGTGACAGTGTGGGGCTGTGGTGGCGCCGCTTACATGACTCG 234
 4799 GGGGAGGAACTGTGGCGGGCGGACCCCGTGTGGCGCCCGCTGACACCGTGTACG 4858
 235 CACACCCAGCAGCAACTGTCTGTGAGCGGTGAGAGAGAGAGAGAGAGAGAGAGAG 294
 4859 GCCCTGACCGGTACCTGTGACCGCGCGCTGCGGAGATCTCCCGCGGGAGAGAGCTG 4918
 295 GCGGTACCAAGGTGTGACAGGGCAATTGGCGCATGAGCTGGCGCGCGCTGGCGGACGCC 354
 4919 GACCTTACCGGTGTACACCGACCGCGCGCTTGTGCGCTTGTGAGTGTGAGTGTGAG 4978
 355 AACCCG 414
 4979 CTGGAACACCAAGCGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5038
 415 GCGGACTACCTGTGACAGCGGTGTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474


```

Sequence 30, Application US/09860846
Patent No. US20020164742A1
GENERAL INFORMATION:
APPLICANT: Sherman, D.H.
APPLICANT: Liu, H.
APPLICANT: Xue, Y.
APPLICANT: Zhao, L.
TITLE OF INVENTION: DNA encoding methymycin and pikromycin
FILE REFERENCE: 600.4380U1
CURRENT APPLICATION NUMBER: US/09/860,846
CURRENT FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: 09/105,537
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 43
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 30
LENGTH: 13842
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-860-846-30

```

Query Match	4.18;	Score 72.2;	DB 9;	Length 13842;
Best Local Similarity	44.88;	Pred. No. 2.5e-05;		
Matches 413; Conservative	0;	Mismatches 493;	Indels 15;	Gaps 3

[illegible]

Db	1003	CTCTCCCGGGGCGCTCGCCGACGAGCCGACCCGGTCTCTGCGCGTATCCGGGGCAGCGCGCTC	1062
OY	673	CATGGCGGCATCTACGGCCACCACCATGCGCGAGAGAGGGAGACACATGCTGGCCGCGCTG	732
Db	1063	AACAAACGGCGCGCGCGCCGACGGGCATGACACCCCGACGCGCGACG-----GCGAG	1113
OY	733	GAGGAACCTTCCGATCGGCGCGCGCGAGCTGGAGAGTGCCGCTGATCTCGCACCAACAG	792
Db	1114	GAGCGCGTGCCTCCGCGAGGCGCCACGAGAGGGGCGCGGACCGCGCGCGCGAGACTGGGTAC	1173
OY	793	GTCATGGGCGCCAGCCCAATTTCGGCGCGCTGCGCGAGACGCTCCGCTGATCGAGGCGCGC	852
Db	1174	GTCAGGCTGACGGCACGGCACCGCCACCCCGTGGGGAGACCCGATGAGAGCGCGCTGCCTCGCG	1233
OY	853	ATGGCGCGCCAGGACGCTCTGCTGGAGCGGGTATCCCTACGTGGCGGGGTCCACCATGCTC	912
Db	1234	GCGCGCCTCGGCACTCGGCCCGCGCGCGGACACCGCTCTGTGGGCTCGGTCAAGCG	1293
OY	913	AAGCAGAGACCGCGTGTCTGCTGGCGGAGCAGCACCATCATCCTGGTGGACGCTTCCCC	972
Db	1294	AACATCGGCCAC---CTGAGAGGCGCGGCGGAGATGCGCGCTCATCAAGCGCGTCTCTG	1350
OY	973	GAACCTGAGCGGGCGGCACCTG	993
Db	1351	GCGGTCCGCGTCCGCGGTG	1371

```

US-09-861-289-30
: Sequence 30, Application US/09861289
: Patent No. US2007020110897A1
: GENERAL INFORMATION:
: APPLICANT: Sherman, D. H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymyc
: FILE REFERENCE: 600.438US1
: CURRENT APPLICATION NUMBER: US/09/861,289
: CURRENT FILING DATE: 2001-05-18
: PRIOR APPLICATION NUMBER: 09/105,537
: PRIOR FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0.
: SEQ ID NO 30
: LENGTH: 13842
: TYPE: DNA
: ORGANISM: Streptomyces venezuelae
: US-09-861-289-30

```

Query Match	4.1%;	Score 72.2;	DB 10;	Length 13842;
Best Local Similarity	44.8%;	Pred. No. 2.5e-05;		
Matches 413; Conservative	0;	Mismatches 493;	Indels 15;	Gaps 3;

QY	76	GGGGGACCCCTCATCGAGGGAGCAACCCCGGGGGGGCGCCGCCCACTGAGCGTGGC	13
Db	463	GTCACACCGGTTGAGAGCGGGCTTTCTTGCGGACTCTGGCCCGCGAGGGCGCGGAGATGGAC	522
QY	136	GGGACCCCGCATCGCGCCCATCGGAGATCTGTGCGAGCGCGCGCGCGCACACCCGGGTGCAC	197
Db	523	CCGAGAGACGCGCTGCGCCCTTGAGACTGTGGGTGGGAGGACCTCGAGGCGCGCGGGATGCAC	582
QY	196	GTTGTCGGGGCCTGGTGTGTCGGCGC---CCGGGCTCATGTCACTGCGACACCCAGAGACAAC	253
Db	583	CCGTCCTCGCTCACCGGGACCCCGACCGGGCGTCTTCCCGCGGCATCTGGAGAGACTTAC	642
QY	253	TACCTGCTCAGGGCCTCGCGACATGACCCGCCAAGATCTGCGAGGGCGTTCACACACGTTGTC	312
Db	643	GCCACCCCTGAAGACCGCGCAAGGGGGGGCGCGCGATCAACCCCGCACACGTCATACCGGCTTC	702
QY	313	ACGGGCAATTGCGGCAATCAGCGCTGGCGCGCGTGGCGAGCGCAACCCCGCGCCCGCCCTTG	372

Db 703 CACCGGCGCATCATCGGAAACGACTCTGTACACGCTCGGGCTCGGGCCAGCATG 762
Qy 373 GACCTGCGGAGGAGGCGGCTTTACCGTTTACGCGCTTCCGCGACTACGTGAGCGC 432
Db 763 GTCTGCACTCCGCGCACTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 822
Qy 433 TTGCGGCGCAACGCGCGGCGGCTGACGCGGCTGTATGTGTGGCCATTACGCTGCGC 492
Db 823 CGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 882
Qy 493 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 552
Db 883 AGCATCTATCGGGGCGGAGGAGTTCGCGGCTCTCTCCCGCGGCGGCGGCGGCGGCGG 942
Qy 553 GACCTGCGGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
Db 943 GACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1002
Qy 613 CCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
Db 1003 CTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1062
Qy 673 CATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 732
Db 1063 AACAAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1113
Qy 733 GAGGAAACCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 792
Db 1114 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1173
Qy 793 GTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852
Db 1174 GTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1233
Qy 853 ATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 912
Db 1234 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1293
Qy 913 AAGCAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 972
Db 1294 AACATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1350
Qy 973 GAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 993
Db 1351 GCGGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1371

RESULT 12

US-09-860-846-5
; Sequence 5, Application US/09860846
; Patent No. US20020164742A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/860.846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: 09/105.537
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-860-846-5

Query Match 4.1%; Score 72.2; DB 9; Length 36778;
Best Local Similarity 44.8%; Pred. No. 2.5e-05;
Matches 413; Conservative 0; Mismatches 493; Indels 15; Gaps 3;

Qy 76 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 135
Db 2204 GTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2263
Qy 136 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195
Db 2264 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2323
Qy 196 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 252
Db 2324 CCGTCTCTGCTACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2383
Qy 253 TACCTGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 312
Db 2384 GCCACCTGAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2443
Qy 313 ACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372
Db 2444 CACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2503
Qy 373 GACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 432
Db 2504 GTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2563
Qy 433 TTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 492
Db 2564 CCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2623
Qy 493 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 552
Db 2624 AGCATCTATCGGGGCGGAGGAGTTCGCGGCTCTCTCCCGCGGCGGCGGCGGCGGCGG 2683
Qy 553 GACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 612
Db 2684 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2743
Qy 613 CCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 672
Db 2744 CTCTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2803
Qy 673 CATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 732
Db 2804 AACAAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2854
Qy 733 GAGGAAACCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 792
Db 2855 GAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2914
Qy 793 GTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 852
Db 2915 GTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2974
Qy 853 ATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 912
Db 2975 GCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3034
Qy 913 AAGCAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 972
Db 3035 AACATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3091
Qy 973 GAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 993
Db 3092 GCGGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 3112

RESULT 13

US-09-861-289-5
; Sequence 5, Application US/09861289
; Patent No. US20020110897A1
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.

Tue Jan 14 09:30:31 2003

us-10-009-782-1.mli

Page 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2003, 06:29:34 ; Search time 74 Seconds
(without alignments)
7285.645 Million cell updates/sec

Title: US-10-009-782-1

Sequence: 1 gaattccattgacgcgga.....ccctgagctacgagaagctt 1758

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCrUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94.2	5.4	8438	1	US-07-945-283-1
2	81.8	4.7	1248	4	US-09-105-537-7
3	81.8	4.7	13613	4	US-09-105-537-3
4	77.4	4.3	23673	4	US-09-773-816-1
5	76.2	4.3	5970	3	US-09-320-878-21
6	75.6	4.3	43280	2	US-08-804-227C-1
7	75	4.3	1320	2	US-08-461-775-8
8	75	4.3	1620	2	US-08-461-775-10
9	75	4.3	1620	2	US-09-031-606-11
10	75	4.3	1620	2	US-09-031-606-10
11	74.4	4.2	12001	1	US-08-458-568A-11
12	73.8	4.2	15872	1	US-09-105-537-1
13	73.4	4.2	2167	2	US-08-461-775-9
14	73.4	4.2	2167	2	US-09-031-606-9
15	73.4	4.2	2668	2	US-08-461-775-11
16	73.4	4.2	2668	2	US-09-031-606-11
17	73.4	4.2	4257	2	US-08-690-473-1
18	73.4	4.2	4257	4	US-09-259-821A-1
19	73.4	4.2	4257	4	US-08-843-659-1
20	72.8	4.1	11604	4	US-09-385-028-13
21	72.8	4.1	15079	4	US-09-385-028-1
22	72.2	4.1	13842	4	US-09-105-537-30
23	72.2	4.1	36778	4	US-09-105-537-5
24	72.2	4.1	38506	3	US-09-320-878-19
25	71.2	4.1	1833	2	US-08-403-852D-6
26	71.2	4.1	1833	3	US-08-510-646B-6
27	71.2	4.1	1833	4	US-09-231-818-6

28	71.2	4.1	2634	3	US-08-941-936-1	Sequence 1, Appl1
29	70	4.0	4403765	4	US-09-103-840A-2	Sequence 2, Appl1
30	69.6	4.0	11220	4	US-09-105-537-32	Sequence 32, Appl1
31	68.8	3.9	5392	2	US-08-403-852D-1	Sequence 1, Appl1
32	68.8	3.9	5392	3	US-08-510-646B-1	Sequence 1, Appl1
33	68.8	3.9	5392	4	US-09-231-818-1	Sequence 1, Appl1
34	68.6	3.9	44377	2	US-08-804-227C-7	Sequence 7, Appl1
35	68.6	3.9	44377	2	US-08-804-198-1	Sequence 1, Appl1
36	68.2	3.9	2220	4	US-08-765-907A-14	Sequence 14, Appl1
37	68.2	3.9	4496	4	US-08-765-907A-6	Sequence 6, Appl1
38	68	3.9	2064	1	US-08-343-428C-1	Sequence 1, Appl1
39	68	3.9	12588	2	US-08-105-537-13	Sequence 13, Appl1
40	67.8	3.8	1014	4	US-08-125-468-1	Sequence 1, Appl1
41	67.4	3.8	30001	1	US-08-474-933-1	Sequence 1, Appl1
42	67.4	3.8	30001	2	US-08-403-852D-4	Sequence 4, Appl1
43	67.2	3.8	1208	3	US-08-510-646B-4	Sequence 4, Appl1
44	67.2	3.8	1208	3	US-09-231-818-4	Sequence 4, Appl1
45	67.2	3.8	1208	4	US-09-231-818-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-07-945-283-1/C
Sequence 1, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudorabies Virus Deletion Mutants
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4128
TELEFAX: 309-685-4011 ext. 513
INFORMATION FOR SEQ ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8438 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pseudorabies virus
FEATURE:
NAME/KEY: CDS
LOCATION: 622..6495
FEATURE:
NAME/KEY: variation
LOCATION: replace(1099, "g")
FEATURE:

QY	212	TCGCGCCGGGCTTCATGCACTGCGACACCCAGACAACTACCTGGCTCAGGGCGTCGG	271
Db	395	CCGCGGGGGTACACACCCCGACTCTGGGCGGTGTGGGCTTCAACCTCTGGGGCGCCCT	454
QY	272	ACATGACGCCCAAGATCTCGACGAGGGGTCACCAAGGTGGTCACGGGCAATTTGGCGCATCA	331
Db	455	GCGCGCGCCGACACAGTGTCCGGAAATGTCGCGAGAGCAGCGCTGGGGGTGTACTTCGACG	514
QY	332	GCCCTGGGGCGGTGCGGCACGCCAACCCGCGCCGCCCTTGGAACTGCTTGGACAAAGCG	391
Db	515	CCGCGCACGGCTCTGGGTGTGGCGGTGTGACGGCGCGCGCCGCGGACCTTCGGCGACGGCG	574
QY	392	GCCTTACCTTCAGTTCGAGCGCTTGGCGGACCTTACCTTGGACGCTTTGGGGGACAGCGGGCG	451
Db	575	AGGTCTTCAGTTGCA---GCGCACCAAGGCGCTCAACCGCTTCGAGGGCGGGCGGCTCG	631
QY	452	CCGTCAACGCCGCGCTGTATGTGTGGGCCATTCAACGCTGCGCGCGCGGTATGCCGGA	511
Db	632	TCACCGACGACGGCGACCTCTGCGCGCCCGGATCCGCCCTCCACAACTTCGGCTTCGACG	691
QY	512	TGCAAGCGGCGCGCCACCGSAGAGAAATGGCGGCCAATGGGGGAACTGGCGGAGAAAGCA	571
Db	692	TGCGCGGGGAGACCCCGCGGGGACCAAGCCCAAGATGACGAGGCGCGCGCGCA	751
QY	572	TGGCAGACGGGCGCATTCGACGATTCGACCGGCGCTTCACCGCGCGCGCGCGCGCGCA	631
Db	752	TGGGCTCTACCTCCCTGTGACGCGCTTTCGGAAGTCTATGACCGGAACGGGGGCAACG	811
QY	632	CCACCGGAAGATATCATGAGAGTGTGCCGGCGGTGAGCGCGCATGGCGGATCTTACGCCA	691
Db	812	CCGCGCTACCGCGAGCACCTTGGCGGACCTCCCGCGCTCTCTGTGCGCGACACGACCCG	871
QY	692	CCCAACATGCGCGAGAGGGGACCATCTGTGCGCGCGGTGGAGAAACCTTCGCATCG	751
Db	872	ACGGGCTCAACAAACACACCATATGATGTGTGCAATTCGACGAGGCGCACCGGGCATTC	931
QY	752	GCCCGGAGCTGGAGCTGCCGGTGGTGATCTCGACCAAGAAGTATGGGCGACGCCAATT	811
Db	932	ACCCCGACACTGTATGAGAGTCTCTGAAGGCGCGAAGGCTGACACACCCGCTACTTCT	991
QY	812	TCGGGCGGCTGGCGAGACGCTCCGCTGATGTCAG-----GCCGCAATGGGGCGCC	862
Db	992	CGCGGGGTGGCAGAGCTGTAGACCGCGGAGCGGAGCGACGCGCCGCTGCGGACCA	1051
QY	863	AGGACGTCTCGCTGAGAGGCTATTCCTACGTGCGCGGCTTCACACATGCTCAAGCAGAGAC	922
Db	1052	CCGAACGCTTCGCGCGCGGTGTCTGTCTCTGCGGACCGGACCGCATCTGGCGACGACG	1111
QY	923	GCGTGTCTGTGGCGGACGACCATCATCTGATGCAAGCCCTTCCCGAACTGAGCG	982
Db	1112	ACATTCGCGCGGGTGGCGGACCTGTGCTGTGCGCGACCGCGGCGCGAACTGACG	1172
QY	983	GGGCGAGCTGATGATGATGGCGGGCGGAGCGCG	1015
Db	1172	CGCGCACCGGACAGGCGCCCGCGCGGTGCG	1204

```

RESULT 3
US-09-105-537-3/C
: Sequence 3, Application US/09105537A
: Patent No. 6265202
: GENERAL INFORMATION:
: APPLICANT: Sherman, D. H.
: APPLICANT: Liu, H.
: APPLICANT: Xue, Y.
: APPLICANT: Zhao, L.
: TITLE OF INVENTION: DNA encoding methymycin and pikromycin
: FILE REFERENCE: 600,438US1
: CURRENT APPLICATION NUMBER: US/09/105,537A
: CURRENT FILING DATE: 1998-06-26
: NUMBER OF SEQ ID NOS: 43
: SOFTWARE: FastSeq for Windows Version 3.0

```

```

; SEQ ID NO 3
; LENGTH: 13613
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-3

```

Query Match	4.7%;	Score 81.8;	DB 4;	Length 13613;
Best Local Similarity	44.2%;	Pred. No. 2.1e-07;		
Matches 439;	Conservative 0;	Mismatches 542;	Indels 12;	Gaps 2;

QY	32	CCAGTGTCCCAATCCGAGATTCCTCCGACCCCTTGAACCTGTGATCTGTGATGGGGGGGAGACCTCATG	91
Db	12659	CCACCTGCAACGCCACGGGCGGGGCTCTCACTGTCTTGGCGACAGCGCGGCTTACCGGGC	12610
QY	92	ACGGCAACAACACCCCGGGGGGAGCGGCGGACGTGGGCTGGCGGGGACCGCATGCGGC	151
Db	12609	AAGGATCATGATCCGTGATGATGATGTTGGCGCCACCCCGCACGACTCGCTGTGATGGCGC	12550
QY	152	CCATCGGGCGATCTGTGTGGAGCGCGCGCGGCGACACACCGGCTGTGAGATGTGGGGCTGTGG	211
Db	12549	TCACCCCGGCTCTTGGCGCACTGTGACCCGGAGACCGGGAACCTTGACCCGGACAGATGG	12490
QY	212	TCGGCCCCGGCTTCATGACTGTGACACACCGACAGACAACTACTGTCTGAGCGCTGGC	271
Db	12489	CCGGCGGGGTACACACCCCGACCTGGGCGGCTGTGGGCGTCACTCTGTGGGGCGCGCCCT	12430
QY	272	ACATGACGCCCAAGATCTGTGGCAAGGCGTCTACACGATGTGTACAGGGCAATTTGGCGATCA	331
Db	12429	GGCGCGCGACACACGTCTGGGAAGTGCGCGAGACACAGGCTTGAGCTGTACTTTGAGAC	12370
QY	332	GCTTGGGCGCGCTGTGGGCGACAGCCACCGCGCGCGCCCTGTGAGCTCTGTGAGAGAGCG	391
Db	12369	CCGGGACAGCGCTCTGGGTGGGTGTGAGGGCGGGCGCGCGGACGCTGTGGGAGCGCGC	12310
QY	392	GCTTGTACCTTTGTGAGCGCTTGTGCGACTACTGTGAGCGTTCGGGGCCACGCGCGG	451
Db	12309	AGGTCTTACCTTCA---CGCACCAAGGCGGTGACAGCTTTGAGGGGGGCGCGCTGTG	12253
QY	452	CCGTCAACGCGGCTGTATGTGTGGGCACTTCAACGCTCGGCGCGCGGTATCGCGGACT	511
Db	12252	TCACGAGCAAGCGCACTGTGCGCGCGCGGATCCGGCGCTTCACAACTTGGGCTTGGACC	12193
QY	512	TGCAAGCGCGCGCACCGAGAGAGAAATCGCGCGCATCGGAGCTTGGCCGAGAGACCA	571
Db	12192	TGCCCGGCGGACCGCGCGCGGGGAGCCAAACGCAATGAGAGAGAGCGCGCGCGCA	12133
QY	572	TGGCGACGGGCGCATCTGGGATTTGTGACCGCGGGCTTGTACCGCGCGCGCGCGCGCA	631
Db	12132	TGGGCGCTCACTCTGTGAGCGCTTTCGCGAGGTATGACCGGAGACCGCGCAACACAG	12073
QY	632	CCACGAGAAGATCATCTGAGGTGTGCGCGGCGCTGAGCGCACATGCGGCACTTACGCCA	691
Db	12072	CCGCGTACCGCGAGACACTGTGCGGACTCCCGGGGCTCGTGTGCGGACACGACCGCG	12013
QY	692	CCCACTATGCGGAGAGAGCGACACATGTGTGGCGCGCTGAGAGAACTTCCGCATG	751
Db	12012	ACGGCTCTAACACACACAGTACTGTATGTGTGAGATGAGACGAGCCACACCGGATCC	11953
QY	752	GCCCGGAGCTGAGAGTGTGCGGTGTATGTCTGCACACAAAGTATGAGCGCACCAATT	811
Db	11952	ACCGGAGCTGTATGTAGAGTCTGTGAAGCGGAGAGGGGTGTACACCGCGGCTACTTCT	11893
QY	812	TCGGCGCTGTGCGGAGAGCGCTGCGCTGATGAG-----GCCGCATGGCGCGCC	862
Db	11892	CGCGGGGCTGTGCAGAGCTGTGAGCCCTTACCGGGGACACCGCACCGCCGCTGTGCGACA	11833
QY	863	AGGAGCTGTGCTGTGAGAGCGCTATCCCTACTGTGGCGCGGTGTACACATGTCTAAAGAGGAC	922
Db	11832	CCGAAACGCTGTGCGCGCGGCGGTCTTCCCTGTCCGACCGGCAACCGCATGTGGCGAGACG	11773
QY	923	GCGTGTCTGTGGCGGACGCGACCATCATCACCTGTGGTGCAAGCCCTTCCCGAATGTAGCG	982
Db	11772	ACATTCGCGGGGTGTGCGGACTCTGTCTGTGTGGCGACACCGGGGCGGAGATGTAGCG	11713

RESULT 5
US-09-320-878-21/c
Sequence 21, Application US/09320878A
Patent No. 6117659
GENERAL INFORMATION:
APPLICANT: ASHLEY, Gary
APPLICANT: BETLACH, Melanie C.
APPLICANT: BETLACH, Mary C.
APPLICANT: MCDANIEL, Robert
APPLICANT: TANG, Li
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 300622002120
CURRENT APPLICATION NUMBER: US/09/320,878A
EARLIER FILING DATE: 1999-05-27
EARLIER APPLICATION NUMBER: CIP OF 09/141,908
EARLIER FILING DATE: 1998-08-28
EARLIER APPLICATION NUMBER: CIP OF 09/073,538
EARLIER FILING DATE: 1998-05-06
EARLIER APPLICATION NUMBER: CIP OF 08/846,247
EARLIER FILING DATE: 1997-04-30
EARLIER APPLICATION NUMBER: 60/119,139
EARLIER FILING DATE: 1999-02-08
EARLIER APPLICATION NUMBER: 60/100,880
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: 60/087,080
EARLIER FILING DATE: 1998-05-28
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 5970
TYPE: DNA
ORGANISM: Streptomyces venezuelae
US-09-320-878-21

Query Match 4.3%; Score 76.2; DB 3; Length 5970;
Best Local Similarity 44.7%; Pred. No. 2.2e-06;
Matches 334; Conservative 1; Mismatches 409; Indels 3; Gaps 1;

0Y 32 CCATGTCCCAATCCGANTCCAGCCCTTCGACCTGCTGCGGGGGGACCTCATCG 91
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 781 CCACCTCGAAGCGACGCGGCTCCAGCTCTCGCGACGCGCGGCTCAACCGCGC 722
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 92 ACGGAGCAACACCGCGGGGGGCGCGCGGACCTGCGGCGGCGGCGGCGGCGGCGG 151
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 721 AAGTGAATCATCCGTCGATGACGTTCCGCGCCACCGCGACGACGACGCTGATGCGCC 662
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 152 CCATCGGCGATCTGTCGACGCGCGCGCGACACCGCGGCTGACGCTGCGGCTGTTGG 211
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 661 TCACCGCGGCTTTCGCGGACATCGACCGCGGACCGGCAACCTCGGACGAGTGG 602
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 212 TCGGGCGCGGCTTCATGACGTCGACACCGCGGACGACGACGACGACGACGACGACG 271
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 601 CCGCGCGGTCACACCGCGGACCTGCGCGGCTGTCGCGGCTGACCTTGGGGCGGCT 542
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 272 ACATGACGCCCAAGATCTCGAGGCGGTACACGAGTGTGTCAGCGGGAATTGCGGCATCA 331
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 541 GCGCGCGGCGGACGAGTGGGAGAGTGGCGCGGACGAGGAGGAGGAGGAGGAGGAGG 482
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 332 GCGTGGCGCGGTCGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 391
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 481 CCGCGCGGCGGCTGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 392 GCTTTACGTTTGAAGCGCTTCGCGGACGCTGACGAGCGTTGGGCGGCAAGCGGCGG 451
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 421 AGGTTTAAAGCTTC---CACGCCACCAAGCGGCTTAAGGCTTGAAGGCGGCGGCGG 365
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 452 CCGTACAGCGCGGCTGTATGTGTGGGCAATTAACGCTGCGGCGGCGGCTCATGCGGAGT 511
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 364 TCACGAGAGCGCGGACGCTGCGCGCGGAGTCGCGGCGGCTTCACAACTTCGCTTCGAC 305
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 512 TGCAGGCGCGGCGGACGAGGAGGAATTCGCGGCGGCTGCGGCGGAGGAGGCA 571
111 111 111 111 111 111 111 111 111 111 111 111 111 111

DB 304 TGCCCGGCGGACCGCGCGCGGCGGAGCACACCGCAAGATGAGCGGAGCGCGCGCGCA 245
0Y 572 TGCCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 631
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 244 TGCGCTACCTCCCTCGACGCGCTTCCGAGGTCATCGACCGGAACCGGCGCAACCGAG 185
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 632 CCACCGAAGATCATGAGGAGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCA 691
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 184 CCGGCTACCGGCGGACCGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 125
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 692 CCACATGCGGCGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 751
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 124 ACGGCTTAACACACCGACGAGTGTGCGGAGTGTGCGGAGTGTGCGGAGTGTGCGGAG 65
111 111 111 111 111 111 111 111 111 111 111 111 111 111
0Y 752 GCGGAGCTGAGCGTCCGCGGCTGCA 778
111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 64 ACGGAGCTGCTGATGAGGTCCTGA 38
111 111 111 111 111 111 111 111 111 111 111 111 111 111

RESULT 6
US-08-804-227C-1
Sequence 1, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuesters, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: A35
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 43280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 816..14234
FEATURE:
NAME/KEY: CDS
LOCATION: 14351..19945
FEATURE:
NAME/KEY: CDS
LOCATION: 20010..31199
FEATURE:
NAME/KEY: CDS
LOCATION: 31232..36067
FEATURE:
NAME/KEY: CDS


```

Db 4525 CGGCCCCGACCCCGCGCGGAGGAGACCCGCGGCCCTTCCCGCGCGCCG 4466
Qy 844 GAGGCCCCCATGGCGGCGGACGCTCTGAGCGCGATCCCTGAGCGCGGCTCC 903
Db 4465 CACGCCACGCGCGCGCGCGCGCGCTGCGAGCTGCGCTGCGCGGCGCG 4406
Qy 904 ACATGCTCAGACGAGCGCGCTGCTGCGCGCGAGACCATCATCATCTGTCAG 963
Db 4405 CTGCTGCTCATG-----CGCTGCGCGGCGAGCTGCGCGCGCGCGCG 4352
Qy 964 CCGTTCGCGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1023
Db 4351 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4292
Qy 1024 AAGTACGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1083
Db 4291 CGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4232
Qy 1084 CCGGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1143
Db 4231 CAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4172
Qy 1144 CACGACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1203
Db 4171 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4112
Qy 1204 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263
Db 4111 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4052
Qy 1264 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1323
Db 4051 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3992
Qy 1324 GGTGCTTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1377
Db 3991 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3932
Qy 1378 GAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1437
Db 3931 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3872
Qy 1438 GCGTTCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497
Db 3871 CGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3812
Qy 1498 CAGCGCTTACATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1541
Db 3811 TACTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3768

```

```

RESULT 12
US-09-105-537-1
; Sequence 1, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D. H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600,438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1 15872
; LENGTH: 15872
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-1

```

```

Query Match 4.2%; Score 73.8; DB 4; Length 15872;
Best Local Similarity 43.2%; Pred. No. 7e-06;
Matches 586; Conservative 0; Mismatches 747; Indels 24; Gaps 4;

Qy 175 GCGGCGACACCGGCGTGCAGCGTGTGCGCGCGCGCGCGCGCGCGCGCG 234
Db 4799 GCGGCGACACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4858
Qy 235 CACACCGACGACGACACTGCTCAGGCGTGCAGACATGAGCGCGCAAGTCTCGAG 294
Db 4859 GCGCTGACCGTACCTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4918
Qy 295 GCGGTACGACGCGTGTGCAGCGCGCAATGGCGCATGACCTGGCGCGCGCG 354
Db 4919 GACCTACCGCGTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 4978
Qy 355 AACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 414
Db 4979 CTGGAACACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5038
Qy 415 GCGGACTACCTGAGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
Db 5039 GCGCGCGACGTGCGCGCGTGTCTCTCCCTGAGGAGCGCGCGCGCGCGCG 5098
Qy 475 GCGCATTCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 534
Db 5099 GCGCGCGCTCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5158
Qy 535 GAATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
Db 5159 GCGGAGGTCTGCGAGTCCCTGMAAGGCTACGAGGCGAGGCTCCCGCGCG 5218
Qy 595 TCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654
Db 5219 GGACCCACCGCGCGCGTGTGCTCCGCGCGAGCGCGCGCGCGCGCGCGCG 5278
Qy 655 TCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714
Db 5279 TGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5338
Qy 715 CACATCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774
Db 5339 CACATGAGCGAGCTGTGAGGAGTCTCTCGGCGCGCGCGCGCGCGCGCG 5398
Qy 775 GTGATCTGCGACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 834
Db 5399 CCGCGGATCCCGCGTGTCTCCACGCGCGCGCGCGCGCGCGCGCGCG 5451
Qy 835 CCGCTGATCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 894
Db 5452 --GCTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5509
Qy 895 GCGGCTCACCAGTCTCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 954
Db 5510 GCGGCTCACCAGTCTCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 5569
Qy 955 TGTGCGAAGCGCTTCCCGAAGTCAAGCGCGCGCGCGCGCGCGCGCG 1014
Db 5570 GTCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5629
Qy 1015 GGCATATCCAGTACGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCG 1074
Db 5630 GACGTACAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5689
Qy 1075 ATGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1134
Db 5690 GGTTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5749
Qy 1135 GCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1194
Db 5750 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5809
Qy 1195 GCGCAC-----TATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1248

```


REFERENCE/DOCKET NUMBER: 010830-035
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2167 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-09-031-606-9

Query Match 4.2%; Score 73.4; DB 3; Length 2167;
 Best Local Similarity 46.7%; Pred. No. 6.8e-06;
 Matches 233; Conservative 0; Mismatches 266; Indels 0; Gaps 0;

QY 297 CGTACCAAGGTGTGACGGCAATTGGCGCATAGCCCTGGCGCCCTGGCGCCGCA 356
 DB 938 CGGCCCAAGGCGCGCAACGTCTATCGACAGAGTTGCGCCCGCCGACCATACCA 997
 QY 357 CCGCGCCCGCCCTGGACCTGTGAGACAGGCGGCTTACCGTTTCGAGCGCTTCCG 416
 DB 998 CGACGGCGTCAACATCCCTGAGGTGAGTGCAGACACCCGTCAGAGAACCTCGGCG 1057
 QY 417 CGACTACCTGACGGCTTGGCGGCACCGCGCGCGCTCAACGCGCGCTGTATGTGG 476
 DB 1058 CCAGCTGCTCAAGAGGTGGCGACCAAGACCAATCGCGGGTGAACGACCAACAC 1117
 QY 477 CCATTCAACGCTGCGCGCGCGCGGTATGCCGACTTGCAGAGCGCGCGCCGACGAG 536
 DB 1118 CGGACCGCTGCTGCGCGCGCGGTGTCGCGAGGGCGCTGCGCAAGTGGCGCGC 1177
 QY 537 AATCGCGCATGCGGACCTGGCGAGAGCCATGCGACGCGCGCGCTATGCGATTTC 596
 DB 1178 CTCGCCGCGCGCGCTGAGAGAGGCGATGAGCGCGCGCGCTGCGCGCTGCGGAG 1237
 QY 597 GACCGCGCTTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 656
 DB 1238 GCTGCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 1297
 QY 657 CGCGCGCGTGAAGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 716
 DB 1298 CGCGCGAGCAAGAGGTGGCGAGCTCATCGCCGAGCGAGTGAACAGGTGGCGAG 1357
 QY 717 CATCTGCGGT 776
 DB 1358 CGGTGTCAATCAAGTGAAGAGTCCCAACACTTCGGTGTGCACTGACCTTACCG 1417
 QY 777 GATCTGCAACCAAGTTC 795
 DB 1418 CATGCGCTTCGACAAAGGCG 1436

RESULT 15
 US-08-461-775-11
 Sequence 11, Application US/08461775
 Patent No. 5858773
 GENERAL INFORMATION:
 APPLICANT: MAZODIER, Philippe
 APPLICANT: GUGLIEMI, Gerard
 TITLE OF INVENTION: REGULATORY NUCLEOTIDE SEQUENCE OF THE
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Burns, Doane, Swecker & Mathis
 STREET: George Mason Bldg., Washington & Prince Sts.
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/461,775
 FILING DATE: 435
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/050,313
 FILING DATE: 10-MAY-1993
 APPLICATION NUMBER: FR 9011186
 FILING DATE: 10-SEP-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Crane-Feury, Sharon E
 REGISTRATION NUMBER: 36,113
 REFERENCE/DOCKET NUMBER: 010830-035
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2668 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-461-775-11

Query Match 4.2%; Score 73.4; DB 2; Length 2668;
 Best Local Similarity 46.7%; Pred. No. 7e-06;
 Matches 233; Conservative 0; Mismatches 266; Indels 0; Gaps 0;

QY 297 CGTACCAAGGTGTGACGGCAATTGGCGCATAGCCCTGGCGCCCGCTGGCGCCGCA 356
 DB 938 CGGCCCAAGGCGCGCAACGTCTATCGACAGAGTTGCGCGCGCCCGACCATACCA 997
 QY 357 CCGCGCCCGCCCTGGACCTGTGAGACAGGCGGCTTACCGTTTCGAGCGCTTCCG 416
 DB 998 CGACGGCGTCAACATCCCTGAGGTGAGTGCAGACACCCGTCAGAGAACCTCGGCG 1057
 QY 417 CGACTACCTGACGGCTTGGCGGCACCGCGCGCGCTCAACGCGCGCTGTATGTGG 476
 DB 1058 CCAGCTGCTCAAGAGGTGGCGACCAAGACCAATCGCGGGTGAACGACCAACAC 1117
 QY 477 CCATTCAACGCTGCGCGCGCGCGGTATGCCGACTTGCAGAGCGCGCGCCGACGAG 536
 DB 1118 CGGACCGCTGCTGCGCGCGCGGTGTCGCGAGGGCGCTGCGCAAGTGGCGCGC 1177
 QY 537 AATCGCGCATGCGGACCTGGCGAGAGCCATGCGACGCGCGCGCGCTATGCGATTTC 596
 DB 1178 CTCGCCGCGCGCGCTGAGAGAGGCGATGAGCGCGCGCGCTGCGCGCTGCGGAG 1237
 QY 597 GACCGCGCTTCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 656
 DB 1238 GCTGCACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 1297
 QY 657 CGCGCGCGTGAAGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 716
 DB 1298 CGCGCGAGCAAGAGGTGGCGAGCTCATCGCCGAGCGAGTGAACAGGTGGCGAG 1357
 QY 717 CATCTGCGGT 776
 DB 1358 CGGTGTCAATCAAGTGAAGAGTCCCAACACTTCGGTGTGCACTGACCTTACCG 1417
 QY 777 GATCTGCAACCAAGTTC 795
 DB 1418 CATGCGCTTCGACAAAGGCG 1436

Search completed: January 14, 2003, 06:36:00
 Job time : 334 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2003, 08:03:39 ; Search time 2587 Seconds
(without alignments)
11005.669 Million cell updates/sec

Title: us-10-009-782-1
Perfect score: 1758
Sequence: 1 gaattcactgacgcgcga.....ccctgacgacgagaacgtt 1758

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

EST:*

1: em_estbba:*
2: em_estbba:*
3: em_estbba:*
4: em_estbba:*
5: em_estbba:*
6: em_estbba:*
7: em_estbba:*
8: em_estbba:*
9: gb_estbba:*
10: gb_estbba:*
11: gb_estbba:*
12: gb_estbba:*
13: gb_estbba:*
14: gb_estbba:*
15: gb_estbba:*
16: gb_estbba:*
17: gb_estbba:*
18: gb_estbba:*
19: gb_estbba:*
20: gb_estbba:*
21: gb_estbba:*
22: gb_estbba:*
23: gb_estbba:*
24: gb_estbba:*
25: gb_estbba:*
26: gb_estbba:*
27: gb_estbba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	4.7	1516	12	BG809984 mgct002xd
2	82.8	4.7	1798	17	AG171124 Pan trogl
3	82	4.7	1798	17	AG171124 Pan trogl
4	78.4	4.5	1542	17	AG032943 Pan trogl
5	76	4.3	1189	17	AG030608 Pan trogl
6	74	4.2	1116	12	BG810038 mgct002xi

7	72.4	4.1	925	17	CNS0091P	AL053013 Drosophi1
8	72.4	4.1	1299	17	AG039481	AG039481 Pan trogl
9	71.8	4.1	932	17	CNS00720	AL066742 Drosophi1
10	71.8	4.1	1341	17	AG030611	AG030611 Pan trogl
11	71.6	4.1	925	17	CNS0091P	AL053013 Drosophi1
12	71.2	4.1	1152	17	AG076818	AG076818 Pan trogl
13	71.2	4.1	1538	17	AG030607	AG030607 Pan trogl
14	71	4.0	1129	17	AG127788	AG127788 Pan trogl
15	71	4.0	1281	12	BG852363	BG852363 1024034A0
16	71	4.0	1288	14	BG678719	BG678719 AGENCOURT
17	70.2	4.0	1288	14	CNS006XK	AL066051 Drosophi1
18	70.2	4.0	1131	17	AG042820	AG042820 Pan trogl
19	69.6	4.0	1625	17	AG043477	AG043477 Pan trogl
20	69.2	3.9	1360	13	BM463178	BM463178 AGENCOURT
21	69	3.9	1160	14	BM926731	BM926731 AGENCOURT
22	68.6	3.9	935	17	CNS006XK	AL066051 Drosophi1
23	68.6	3.9	1364	14	BM810045	BM810045 AGENCOURT
24	68.4	3.9	932	17	CNS00720	AL066742 Drosophi1
25	68.4	3.9	982	14	BO887717	BO887717 AGENCOURT
26	68.4	3.9	1626	10	AW731212	AW731212 GA_Ea001
27	68.2	3.9	1387	10	AW731151	AW731151 GA_Ea001
28	67.8	3.9	936	12	BG852371	BG852371 1024034A0
29	67.8	3.9	1232	17	AG072425	AG072425 Pan trogl
30	67.8	3.9	1328	17	AG043615	AG043615 Pan trogl
31	67.6	3.8	1088	17	AG073424	AG073424 Pan trogl
32	67.6	3.8	1299	17	AG039481	AG039481 Pan trogl
33	67.4	3.8	1189	17	AG162326	AG162326 Pan trogl
34	67.4	3.8	1893	17	AG030608	AG030608 Pan trogl
35	67.4	3.8	1278	17	AG060116	AG060116 Pan trogl
36	67.4	3.8	1651	12	BG809816	BG809816 mgct001xk
37	67.2	3.8	644	12	BG607300	BG607300 WHE2493_G
38	67.2	3.8	1248	17	AG049208	AG049208 Pan trogl
39	66.6	3.8	1040	12	BG786331	BG786331 SEAMC006
40	66.4	3.8	1366	17	AG065058	AG065058 Pan trogl
41	66.2	3.8	1086	17	AG077581	AG077581 Pan trogl
42	66.2	3.8	1101	17	AG039543	AG039543 Pan trogl
43	66	3.8	1164	17	AG128373	AG128373 Pan trogl
44	65.8	3.7	1339	13	BM458211	BM458211 AGENCOURT
45	65.8	3.7	1452	17	AG032979	AG032979 Pan trogl

ALIGNMENTS

RESULT 1
LOCUS BG809984 1516 bp mRNA linear EST 22-MAY-2001
DEFINITION mgct002xd11f Magnaporthe grisea Appressorium Stage cdna Magnaporthe
grisea CDNA clone mgct002xd11f 5', mRNA sequence:
BG809984
VERSION BG809984.1 GI:14180964
KEYWORDS EST.
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea.

REFERENCE
AUTHORS Sordariomyces Incertae sedis; Magnaportheaceae; Magnaporthe.
TITLE Chol.W. and Dean.R.A.
JOURNAL library in the rice blast fungus, Magnaporthe grisea
COMMENT unpublished (2001)

CONTACT: Ralph A. Dean
Fungal Genomics Laboratory
North Carolina State University
Campus Box 7251, Raleigh, NC 27695, USA
Tel: 919-513-0020
Fax: 919-513-0024

Email: ralph.dean@ncsu.edu
Seq primer: T3 primer (AATTAACCTCCTCAAGCG).
Location/Qualifiers 1. 1516

FEATURES
source /organism="Magnaporthe grisea"
/strain="70-15"

[illegible][illegible]

RESULT 2	1798 bp	DNA	linear	GSS 09-JAN-2002
AG171124	pan troglodytes	DNA, clone: RP43-040F09.TJ,	genomic survey	
LOCUS	sequence.			
DEFINITION	AG171124			
ACCESSION	AG171124			
VERSION	AG171124.1	GI:16700802		
KEYWORDS	GSS.			
SOURCE	pan troglodytes	male lymphocytes	DNA, clone: RPCI-43	Chimpanzee

ORGANISM	Pan troglodytes
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE	BAC end sequences of library RPCR-43
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1798)
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (02-AUG-2001), Asao Fujiyama, The Institute of Physical

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chiampesegsc.riken.go.jp, URL: <http://hnp.gsc.riken.go.jp/>,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
COMMENT end was generated during the Rad process and may have higher chances
of clone tracking errors.

```

Vector      : PBACE3.6
R.Site 1   : ECORI
R.Site 2   : ECORI
Location/Qualifiers
1. 1798
/organism="Pan troglodytes"
/db.xref="taxon:9598"

```


[illegible][illegible]

QY 1337 CGCCAGGATGCGCATACCGACCTTGAACACCTACGAGCGCGCGGACATCC 1396
 DB 427 GGGCCNNCCCCCCCCCCCCNNCCNNCCCCCCCCCCCCCCCCCCCCNNC 368
 QY 1397 ATTCGCTACGTCACAGCGCGCGGCTCTGCAAGACAGCGCTTACCGGACGATG 1456
 DB 367 GGGCGCCCG 308
 QY 1457 CGCGCGCGGCTGCTGCGACGAGCGCGCGCTGAGCGCGCGCGCGCGCGCGCG 1516
 DB 307 CG 248
 QY 1517 TGAACGGGCTC 1576
 DB 247 CCCCCCG 188
 QY 1577 CCGCGCTGCGCATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1602
 DB 187 NNNNNCCCG 162

RESULT 7

CNS0091P 925 bp DNA linear GSS 03-JUN-1999
 LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
 DEFINITION BAC19D16 of RPCI-98 library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

GSS.
 Drosophila melanogaster.
 Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 925)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : sequefgenoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see <http://www.fruitfly.org> The BDGP Drosophila
 melanogaster BAC library was prepared by Kazutoyo Osoegawa and
 Aaron Mamoser in Peter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
 1. 925
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="BAC19D16"
 /clone_lib="RPCI-98"
 /note="end : TET3"
 BASE COUNT 120 a 61 c 61 g 172 t 511 others
 ORIGIN

Query Match 4.1% Score 72.4; DB 17; Length 925;
 Best Local Similarity 13.3%; Pred. No. 0.017;
 Matches 51; Conservative 190; Mismatches 143; Indels 0; Gaps 0;
 QY 1092 GGAGGCGATCCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1151
 DB 542 GGTTCGTTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 601

QY 1152 CGCCCGCATCCGCGCTGCGGCGACCTTCCCGGCGCTCTGGGCGCATATCCGCGCA 1211
 DB 602 SSSSSSSKCSSTSBSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSCSC 661
 QY 1212 CCTGGCGCTGTCGCGGCGAGCGCGCGGATAGCATGACGCGCGCTGACCCCGCGCG 1271
 DB 662 SSSSSSTTTKSTBSAGSGSMAGGSGSTGSTSSSSSSSTSTSSSSSSSSSS 721
 QY 1272 CTTCGGCTG 1331
 DB 722 SSSSSSSGSSSSSTSBSCSTSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 781
 QY 1332 CGACCG 1391
 DB 782 GSTGSSSSSVGTSSSDSTSTCCSCCYMCTCTBMBCTYSTGCGSSSSGKGV 841
 QY 1392 CATCATTCGCTACGTCACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCA 1451
 DB 842 KCGCGCGCGSSSTNGBGTSACSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 901
 QY 1452 GCATCCGCGCGCGCGCTGCGCACG 1475
 DB 902 SAKRSSSGSVSGSGSGSGSVSG 925

RESULT 8

AG039481 1299 bp DNA linear GSS 01-NOV-2001
 LOCUS Pan troglodytes DNA, clone: PTB-016J15.F, genomic survey sequence.
 DEFINITION AG039481
 ACCESSION AG039481.1 GI:16568206
 VERSION
 KEYWORDS

SOURCE

ORGANISM

Pan troglodytes male lymphoblast DNA, clone_lib: PTB Chimpanzee Male
 BAC library clone: PTB-016J15.F.
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbases@sc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the Rsd process and may have higher chance of
 clone tracking errors.

COMMENT

PRIMERS
 Sequencing: -21M13

LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1. 1299
 /organism="Pan troglodytes"
 /db_xref="taxon:9596"
 /clone="PTB-016J15.F"
 /sex="male"
 /cell_type="Lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 24 a 571 c 592 g 31 t 81 others
 ORIGIN

Query Match 4.1% Score 72.4; DB 17; Length 1299;


```

ACCESSION   AC030611
VERSION     AC030611.1
KEYWORDS    GI:16557484
SOURCE      Pan troglodytes male lymphoblast DNA, clone_11b.PTB Chimpanzee Male
            BAC Library clone:PTB-003A20.F.
ORGANISM    Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE   1
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       BAC end sequences of library PTB
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 1341)
AUTHORS     Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       Direct Submission
JOURNAL     Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC):
            1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
            E-mail:chimbdes@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Clones are derived from the chimpanzee BAC library PTB This BAC end
            was generated during the Rtd process and may have higher chance of
            clone tracking errors.
COMMENT     PRIMERS
            Sequencing: -21M13
LIBRARY     Vector : PKS145
            R.Site 1 : SacI
            R.Site 2 : SacI.
FEATURES    location/Qualifiers
            1..1341
            /organism="Pan troglodytes"
            /db_xref="taxon:9596"
            /clone="PTB-003A20.F"
            /sex="male"
            /cell_type="Lymphoblast"
            /clone_11b="PTB Chimpanzee Male BAC Library"
BASE COUNT  48 a 476 c 700 g 14 t 103 others
ORIGIN
Query Match      4.1%; Score 71.8; DB 17; Length 1341;
Best Local Similarity 44.1%; Pred. No. 0.023;
Matches 245; Conservative 0; Mismatches 307; Indels 3; Gaps 1;
QY 1105 GCGTTGGCCCGACCAAGATGAGTCCGACGCGCTCCGCGACGAGACGCGCCGATCCG 1164
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1165 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1106
QY 1165 GCGCTGTGGGGACCTTCCCGCGCGCGCTGTGGGACATATGGCGCGACCTGGCTGTTC 1224
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1105 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1046
QY 1225 CCGGTGAGAGAGCGGCGGTATGAGATGACCGGCTGACCGCGCGCTTGTGGCTGGCC 1284
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 1045 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 986
QY 1285 GGGGCGGGGACGTGACGAGCGGGGTACTTGCAGCGACTGTGTGTTCGACCGCGCACG 1344
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 985 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 926
QY 1345 GTGGCGCATACCGGACCTTGGACACCTTACCGAGCGCGCGCGCGCGCGGATCCATTCG 1404
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 925 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 866
QY 1405 TACGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1461
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 865 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 806
QY 1462 GCGGTGCTGACGACGAGCGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1521
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 805 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 746

```

```

QY 1522 GGGGCGGGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1581
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 745 NMGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 686
QY 1582 GCGGCGAATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1641
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 685 CCGNCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 626
QY 1642 GCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1656
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 625 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 611
RESULT 11
CNS0091P/c
LOCUS
DEFINITION  Drosophila melanogaster genome survey sequence TEP3 end of BAC #
            BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL053013
VERSION     AL053013.1
KEYWORDS    GI:4934461
SOURCE      Drosophila melanogaster.
ORGANISM    Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 925)
AUTHORS     Genoscope.
TITLE       Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
JOURNAL     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutyo Osoegawa and
            Aaron Mammosser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES    location/Qualifiers
            1..925
            /organism="Drosophila melanogaster"
            /db_xref="taxon:7227"
            /clone="BACR19D16"
            /clone_11b="RPCI-98"
            /note="end : TEP3"
BASE COUNT  120 a 61 c 172 t 511 others
ORIGIN
Query Match      4.1%; Score 71.6; DB 17; Length 925;
Best Local Similarity 14.0%; Pred. No. 0.023;
Matches 55; Conservative 189; Mismatches 150; Indels 0; Gaps 0;
QY 1134 GCGCTCCCGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1193
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 925 CSBSCSCSCSSBSSTSSMSBSSSSBSSSTSSMSBSSSSBSSSSBSSSSSSSSSS 866
QY 1194 GGGGCACTATGCGGCGACCTGCGGCGCTTCCCGTGGAGACGCGGATGAGATGAC 1253
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 865 SSGTSSACVACNMASSSGCGCGGACGACGCGCGCGCGCGCGCGCGCGCGCGCG 806
QY 1254 CCGCTTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1313
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

OY	165	GTCGGACCGCCGCGGCACACCCGGGTGAGCTGTGAGGCTTGTTGTCGCCGCTT	224
Db	449	GCCCCCGCCGCCCACCCGNCGCGNNNGGCGCCGNGMNCCTCCGGGGCGCGCCGAG	508
OY	225	CATGACTCGCACACCACGACGACAACACTACTCTCCTCAAGGCGTGGCGCATGAGCCCAA	284
Db	509	GGGCGGAGGGCGCCCGCCGCGCGCGC--CCGGGNNGCCGCGCGCGGNGNGGGGGGGGG	566
OY	285	GATCTCGCAGGGCGCTCAACCAAGGTGCTACAGGGCAATTGGCGCATCAGCTGAGCGCT	344
Db	567	GGGGGCGCGNCCCGC	626
OY	345	GGCGCAGCGCCAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT	404
Db	627	GGC	686
OY	405	CGACCGCTTCCCGCACTACTGACGCGTTTCGG-CCGACGCGCGCGCGCGCTCAACCGC	463
Db	687	CGC	746
OY	464	CTGTATGCTGGGCCATTCAACGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	523
Db	747	GC	806
OY	524	CCACCGACGAGAATAATGCGCGCATGCGGAGCATGCGCGAGGAAGCATGCGCAAGCGCG	583
Db	807	CGNCGC	866
OY	584	CCATCGGCAATTTGACCGCGCGCTTTTACCGCGCGCGCGCGCGCGCGCGCGCGCAAGAGA	643
Db	867	GC	925
OY	644	TGATCGAGCTGTGC	703
Db	926	GC	985
OY	704	ACGAAGCGGACACATCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGCTGG	763
Db	986	GC	1045
OY	764	ACGTGCGCGGTGATCTCG	783
Db	1046	CCCGCGCGCGCGCGCGCGCGCG	1065

RESULT 13
AG030607
LOCUS AG030607 1538 bp DNA linear GSS 01-NOV-2001
DEFINITION Pan troglodytes DNM, clone: PTB-003A18.F, genomic survey sequence.
ACCESSION AG030607
VERSION AG030607.1 GI:16557480
KEYWORDS GSS.

SOURCE
ORGANISM Pan troglodytes male lymphoblast DNM, clone_1lb:PTB Chimpanzee Male
Pan troglodytes
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

REFERENCE
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Matsubara,H. and Sakaki,Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
RECORDING 2 (bases 1 to 1538)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Matsubara,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbesegsc.riken.go.jp, URL:http://ngp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the Red process and may have higher chance of

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 14, 2003, 05:26:22 ; Search time 384 Seconds
(without alignments)
10309.930 Million cell updates/sec

Title: US-10-009-782-1
Perfect score: 1758
Sequence: 1 gaattccacttgatcgcgaa.....ccctgagctacgagaagctt 1758

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478
```

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing:	Minimum Match 0%
	Maximum Match 100%
	Listing first 45 summaries

Database :
N.Geneseq_1010192.*
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
1	1758	100.0	1758	22	AAIC91797	Alcaligenes xyloso	
2	1440.6	81.9	1473	21	AAAI51440	DNA encoding an an	
3	103	5.9	114955	20	AAE53491	Human adenosine A	
4	101.2	5.8	312	24	ABN24790	Human ORFX polynu	
5	99.2	5.6	65140	22	AAAD17184	Streptomyces nours	
6	99.2	5.6	125401	22	AAAD17186	Streptomyces nours	
7	94.2	5.4	8438	15	AAO73500	DNA encoding Pseud	
8	93.8	5.3	114955	20	AAE53491	Human adenosine A	
9	93.4	5.3	27541	22	AAAD17185	Streptomyces nours	

10	86	4.9	88421	22	AAAL06781	88421nt genomic DNA
11	84.6	4.8	109519	24	AA508693	Micromonospora DSM
12	81.8	4.7	1248	21	AA287286	S. venezuelae deso
13	81.8	4.7	1248	21	AA287286	S. venezuelae deso
14	81.8	4.7	12441	21	AA287319	S. venezuelae deso
15	81.8	4.7	13613	21	AA287319	S. venezuelae deso
16	81.8	4.7	13613	24	AA287319	S. venezuelae deso
17	79.8	4.5	1320	22	AA022481	grov-1 gene partit
18	78.2	4.4	3957	22	AA096686	HSV-2 immediate ea
19	78.2	4.4	12249	21	AA055840	Complete Mitomyxin
20	78.2	4.4	18331	21	AA055857	Complete Mitomyxin
21	78.2	4.4	154746	24	AA025519	Human herpesvirus
22	78.2	4.4	154746	24	AA025519	Human herpesvirus
23	76.2	4.3	5970	21	AA075635	Nucleotide sequenc
24	76.2	4.3	5970	21	AA075603	Contig 002 from CC
25	76	4.3	12001	16	AA076213	HSV 1/LST region.
26	76	4.3	28598	17	AA076769	Sorangium cellulos
27	76	4.3	58857	21	AA058471	Nucleotide sequenc
28	75.8	4.3	4020	18	AA0791361	Orf virus genomic
29	75.6	4.3	43280	10	AA050413	Tyrosine synthase
30	75.4	4.3	2061	10	AA022408	Sequence encoding
31	75	4.3	1620	13	AA022482	grov-1 gene codin
32	75	4.3	2167	13	AA022484	grov-1 gene codin
33	75	4.3	2668	13	AA022485	grov-1 gene. Str
34	73.8	4.2	15872	18	AA078715	Streptomyces venez
35	73.8	4.2	15872	21	AA0787283	S. venezuelae vep
36	73.4	4.2	4257	19	AA068520	The nucleotide seq
37	73.4	4.2	4257	19	AA070362	Infected cell prote
38	73.4	4.2	5224	22	AA030079	L05390 cDNA clone.
39	73.2	4.2	11604	20	AA0232022	Human MTH1 relate
40	72.8	4.1	15204	22	AA014501	Streptomyces clavus
41	72.8	4.1	15079	16	AA0931480	S. clavuligerus cl
42	72.8	4.1	15079	22	AA0144599	Streptomyces clavus
43	72.4	4.1	13189	23	AA051474	Pseudomonas aerugi
44	72.2	4.1	13842	21	AA078297	S. venezuelae mact
45	72.2	4.1	36778	21	AA078318	S. venezuelae pik

ALIGNMENTS

XX	RESULT 1
XX	AAC91797
ID	AAC91797 standard; DNA; 1756 BP.
XX	
AC	AAC91797;
XX	
DT	27-MAR-2001 (first entry)
DE	
XX	Alcaligenes xylosoxidans subspecies xylosoxidans D-aminocyclase DNA
XX	
KW	D-aminocyclase; zinc tolerant host; recombinant production;
KW	zinc enhanced expression; D-form amino acid synthesis;
KW	antibiotic production; peptide drug; pharmaceutical manufacturing;
ds.	
XX	
XX	
OS	Alcaligenes xylosoxidans.
PN	WO200078926-A1.
XX	
PD	28-DEC-2000.
XX	
PF	15-JUN-2000; 2000MO-JP03932.
XX	
PR	17-JUN-1999; 99JP-0170555.
XX	
PA	(AMANO-) AMANO ENZYME INC.
XX	
PI	Takeuchi K, Koide Y, Hirose Y, Moriguchi M, Isobe K.
XX	
DR	WPI; 2001-080828/09.
XX	
DR	P-PSDB; AAB48975.
XX	

Transformed microorganism from zinc-tolerant host for selective
production of D-aminoacylase, useful in synthesis of high
optical-purity D-amino-acids for antibiotic side-chains and peptide
drugs -

Claim 2: Page 14-18; 22pp; Japanese.

The invention relates to a recombinant zinc-tolerant microorganism which
expresses the D-aminoacylase from Alcaligenes xylosoxidans subspecies
xylosoxidans. The presence of zinc ions in the culture medium enhances
expression of the D-aminoacylase gene, and the invention also relates to
the process of recombinantly producing the D-aminoacylase using the
microorganism of the invention. The recombinant microorganism is used
for the selective production of D-aminoacylase, which is useful in the
synthesis of high optical-purity D-form amino acids for use in the
production of antibiotics and peptide drugs. The present sequence
represents DNA encoding D-aminoacylase from Alcaligenes xylosoxidans
subspecies xylosoxidans.

Sequence 1758 BP; 281 A; 662 C; 568 G; 247 T; 0 other;

Query Match 100.0%; Score 1758; DB 22; Length 1758;
Best Local Similarity 100.0%; Pred. No. 5.8e-267;
Matches 1758; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Oy 1 GAATTCACCTTATGCGGAGAGAGATTTCCATGTCCTCCATCCGATTCACGCTTC 60
Db 1 GAATTCACCTTATGCGGAGAGAGATTTCCATGTCCTCCATCCGATTCACGCTTC 60
Oy 61 GACCTGCTCTCGCGGGGCGGACCCCTCATGACGCGACAGCACCCCGGGGCGCGCC 120
Db 61 GACCTGCTCTCGCGGGGCGGACCCCTCATGACGCGACAGCACCCCGGGGCGCGCC 120
Oy 121 GACCTGCGGCGGCGGCGGACCGCATCGCGGCGATCGTGGAGCGCGCGCGCG 180
Db 121 GACCTGCGGCGGCGGCGGACCGCATCGCGGCGATCGTGGAGCGCGCGCGCG 180
Oy 181 CACACCGGGGTCGACGCTGTCGCGGCTGTGTGCGCGCCCGCTTATCAGTCGACACC 240
Db 181 CACACCGGGGTCGACGCTGTCGCGGCTGTGTGCGCGCCCGCTTATCAGTCGACACC 240
Oy 241 CACGACGACACACTACCTGCTCAGGCGTCGCGACATGACGCCCAAGATCTCGAGGGCGTC 300
Db 241 CACGACGACACACTACCTGCTCAGGCGTCGCGACATGACGCCCAAGATCTCGAGGGCGTC 300
Oy 301 ACCAGGTTGTCAGGGGCAATTCGCGATCAGCGTGGCGCGGTCGCGACGCAACCGC 360
Db 301 ACCAGGTTGTCAGGGGCAATTCGCGATCAGCGTGGCGCGGTCGCGACGCAACCGC 360
Oy 361 CCCGCCCCCTGACCTGCTGACGAGAGCGGCTTTACCGTTTGAAGCGCTTCCGCGAC 420
Db 361 CCCGCCCCCTGACCTGCTGACGAGAGCGGCTTTACCGTTTGAAGCGCTTCCGCGAC 420
Oy 421 TACCTGAGAGCGGTTGCGGGCCAGCGCGGGCGGCTCAAGCGCGCTGTATGTGGGCGCAT 480
Db 421 TACCTGAGAGCGGTTGCGGGCCAGCGCGGGCGGCTCAAGCGCGCTGTATGTGGGCGCAT 480
Oy 481 TCAACGCTGGCGCGCGGCTCATGCGGACTTTCAGCGCGCGCGCCACGACGAGAAATC 540
Db 481 TCAACGCTGGCGCGCGGCTCATGCGGACTTTCAGCGCGCGCGCCACGACGAGAAATC 540
Oy 541 GCGGACATGCGGAGACCTGCGGAGAGAACCATGCGACGCGCGCATGCGCATTTTGACC 600
Db 541 GCGGACATGCGGAGACCTGCGGAGAGAACCATGCGACGCGCGCATGCGCATTTTGACC 600
Oy 601 GGGGCTTTCATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 601 GGGGCTTTCATACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Oy 661 CCGCTGAGCGCGCATGCGCGCATCTAGCCACCCACATGCGGACGAGGCGAGCATC 720
Db 661 CCGCTGAGCGCGCATGCGCGCATCTAGCCACCCACATGCGGACGAGGCGAGCATC 720
```

```

Oy 721 GTGGCCGCGCTGAGAGAAACCTTCCGATCGGCGCGAGCTGAGACGTGCCGGTGTATC 780
Db 721 GTGGCCGCGCTGAGAGAAACCTTCCGATCGGCGCGAGCTGAGACGTGCCGGTGTATC 780
Oy 781 TCGCACCACAAAGTATGAGCGGCGACCCCAATTTGGCGCGCTCGCGGACGCTCGCGCTG 840
Db 781 TCGCACCACAAAGTATGAGCGGCGACCCCAATTTGGCGCGCTCGCGGACGCTCGCGCTG 840
Oy 841 ATGAGCGCGCGCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 ATGAGCGCGCGCATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Oy 901 TCCACATGCTCAAGACGAGACCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Db 901 TCCACATGCTCAAGACGAGACCGGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960
Oy 961 AAGCCCTTCCCGGACGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db 961 AAGCCCTTCCCGGACGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Oy 1021 TCCAGTACGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db 1021 TCCAGTACGAGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Oy 1081 GAACCCGACGTCAGCGCGCATCTGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db 1081 GAACCCGACGTCAGCGCGCATCTGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Oy 1141 CCGCAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Db 1141 CCGCAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Oy 1201 TATGCGCGCGCATGCGGCGCTTCCCGCTGAGAGAGCGGCGATGAGATGACCGCGCTG 1260
Db 1201 TATGCGCGCGCATGCGGCGCTTCCCGCTGAGAGAGCGGCGATGAGATGACCGCGCTG 1260
Oy 1261 ACCGCGCGCGCTTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Db 1261 ACCGCGCGCGCTTGGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1320
Oy 1321 CTGGTGTGTTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Db 1321 CTGGTGTGTTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
Oy 1381 CGCGCGCGCGCATCTCGTGTACGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Db 1381 CGCGCGCGCGCATCTCGTGTACGTCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1440
Oy 1441 TTCACCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Db 1441 TTCACCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1500
Oy 1501 CCCTTACATCGCGGCGTGAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
Db 1501 CCCTTACATCGCGGCGTGAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1560
Oy 1561 CGCTACATGCGCGCTTCCCTCGCTGCAATAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1620
Db 1561 CGCTACATGCGCGCTTCCCTCGCTGCAATAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1620
Oy 1621 GTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
Db 1621 GTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1680
Oy 1681 TCGCAGCGCGCGCGCATGCGGCTGCGACCGCTGTCAAGCGCGCGCGCGCGCGCGCG 1740
Db 1681 TCGCAGCGCGCGCGCATGCGGCTGCGACCGCTGTCAAGCGCGCGCGCGCGCGCGCG 1740
Oy 1741 CTGAGCTACGAGAGCTT 1758
Db 1741 CTGAGCTACGAGAGCTT 1758
```


[illegible]

29-MAY-2001; 2001WO-US10836.
30-MAY-2000; 2000US-206132P.
29-AUG-2000; 2000US-228716P.
(CURA-) CURAGEN CORP.
Shinkets RA, Leach MD:
WPI: 2002-106308/14.
P-PSDB; ABP09038.
Novel human polypeptides and polynucleotides useful for diagnosing, preventing and treating cardiovascular disease, neurodegenerative, hyperproliferative disorders and autoimmune disorders
Disclosure; SEQ ID 18057; 1037pp; English.
The present invention describes substantially purified human proteins (referred to as open reading frame, ORF, where X is 1-11491 (see Table 1 in the specification). ABN15762 to ABN27252 encode the human ORF proteins given in ABP00010 to ABP11500. ORF proteins are useful for treating or preventing a pathology associated with an ORF-associated disorder in humans, and in the manufacture of a medicament for treating a syndrome associated with ORF-associated disorder. ORF polynucleotide sequences can be used in gene therapy. ORF sequences can be used in the treatment of cancer. hyperproliferative disorders, cirrhosis of liver, psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage, osteoarthritis, neurodegenerative disorders, disorders related to organ transplantation, cardiovascular diseases, diabetes mellitus, systemic lupus erythematosus, hypertension, hypothyroidism, cholesterol ester storage disease, various immune deficiencies and disorders, infectious diseases, autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. ORF proteins are also useful for treating burns, incisions, ulcers, for treating osteoporosis, bone degenerative disorders, or periodontal disease, and for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues and conditions resulting from systemic cytokine damage.
N.B. The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPo at: ftp://wipo.int/pub/published_pcc-sequences.

sequence 312 bp; 62 A; 92 C; 105 G; 52 T; 1 other;

Query Match	5.88;	Score 101.2;	DB 24;	length 312;
Best Local Similarity	58.58;	Pred. No. 1.6e-07;		
Matches 175; Conservative	0;	Mismatches 124;	Indels 0;	Gaps 0;

	RESULT 4
ID	ABN24790
AB	ABN24790 standard; cDNA, 312 BP.
XX	
AC	ABN24790;
XX	
DT	24-JUN-2002 (first entry)
XX	
DE	Human ORFX polynucleotide sequence SEQ ID NO:18057.
XX	
KM	Human; open reading frame; ORFX gene therapy; cancer; cirrhosis; hyperproliferative disorder; psoriasis; benign tumour; haemorrhage; degenerative disorder; osteoarthritis; neurodegenerative disorder; cardiovascular disease; diabetes mellitus; systemic lupus erythematosus; hypertension; hypothyroidism; cholesterol ester storage disease; immune deficiency; immune disorder; infectious disease; autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis; myasthenia gravis; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	M0200192523-A2.
XX	
PD	06-DEC-2001.

QY	629	CCACCAACCCGAAGAGATCATCGAGAGTGTCCCGCCGCTGACGCGCATATGCGGCATCTACG	688
Dd	2	CGACCAACCCGAAGAGATGATGGGCCCTGGCGGAGCCGCTGACCGAAGCTGGCGGATCTACA	61
QY	689	CCACCCACATGGCGGACGACGAAGCGACGACATCGGCGCCGCGCGAGAGAAACCTTCGCCGA	744
Dd	62	CGACGATCTCGCCACCGAATATGCGCGGATCTTCGATGCGATGAGACGAGGGCTACCCCG	122
QY	749	TGCGCCGCGAGCTGCAAGCTGCCGGTGATCTTCGCACCAAGATCATGAGGCCACGCCA	806
Dd	122	TGCGCAAAAGATCGCGGCGTGGCCGGTGATCTTCACACCTGAATATCCGAGGCCCTGCCA	181
QY	809	ATTTCGGCCGCTCGCGCGAGACGCTGCCGCTGATCGAGCGCGCCATGAGCGCCACAGAGAG	866
Dd	182	ATTGGGGGCGCGGGCGGAGGTGCTGGAAATCCATCGAAGGTGTACACAGTCTATACGCCGG	24
QY	869	TCTTCGCTGGAACCGATATCCCTATAGTGGCCGGCTTCACACATGCTCAAGCAGAGACCCCGTGG	927
Dd	242	TGGGTTCGCACTGCTATCTCCGTACACCGCAACGCTGTCCACGCTGGACCTCAAGCAGAGGTG	300

AA017184
ID AAD17184 standard; DNA: 65140 BP.
AC AAD17184;
XX 29-NOV-2001 (first entry)
XX Streptomyces noursei nys1 DNA of nystatin PKS gene cluster.
XX Polyketide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW antifungal; antibiotic; nys1; ds.
OS Streptomyces noursei.
XX Location/Qualifiers
FH complement (1..1035)
FT /tag- a
FT /product- "NysD2 partial protein"
FT /note- "CDS does not include stop codon"
FT complement (1056..2576)
FT /tag- b
FT /product- "NysD1 protein"
FT 2806..6906
FT /tag- c
FT /product- "NysA protein"
FT 6932..16530
FT /tag- d
FT /product- "NysB protein"
FT 16550..49840
FT /tag- e
FT /product- "NysC protein"
FT 50260..51015
FT /tag- f
FT /product- "NysE protein"
FT 51405..54305
FT /tag- g
FT /product- "NysR1 protein"
FT 54329..57190
FT /tag- h
FT /product- "NysR2 protein"
FT /note- "CDS does not include start codon"
FT 57180..59963
FT /tag- i
FT /product- "NysR3 protein"
FT 60415..61047
FT /tag- j
FT /product- "NysR4 (short) protein"
FT /note- "CDS does not include start codon"
FT 61736..62497
FT /tag- k
FT /product- "NysR5 protein"
FT /note- "CDS does not include start codon"
FT complement (62551..63615)
FT /tag- l
FT /product- "ORF2 protein"
FT /note- "CDS does not include start codon"
FT 63765..64961
FT /tag- m
FT /product- "ORF1 protein"
XX MO200159126-A2.
XX 16-AUG-2001.
XX PD 08-FEB-2001; 2001MO-GB00509.
XX PF 08-FEB-2000; 2000GB-0002840.
XX 08-FEB-2000; 2000GB-0002840.
PR 10-APR-2000; 2000GB-0008786.
PR 14-APR-2000; 2000GB-0009387.
XX (UNO-) UNIV NORGES TEKNISK NATURVITENSKAPELIGE.
PA (SWF) SINTER STIFTELSEN IND TEK FORSK.
PA (ALPH-) ALPHARMA AS.

PA (SINV-) SINVENT AS.
PA (DZIE/) DZIEGEMSKA H.
PA (ZOTC/) ZOTCHEV S.B.
PA (SEK/) SEKUROVA O.N.
PA (FJAE/) FJAEVRIK E.
PA (BRAU/) BRAUTASET T.
PA (STRO/) STROM A.R.
XX Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
PI Valla S, Ellingsen TE, Sletta H, Gulliksen O;
XX WPI: 2001-557614/62.
DR P-PSDB: AAE10125, AAE10126, AAE10127, AAE10128, AAE10129, AAE10130,
XX AAE10131, AAE10132, AAE10133, AAE10134, AAE10135, AAE10136, AAE10137.
XX New nystatin polyketide synthase polynucleotides and polypeptides,
XX useful as antibiotics and antifungals -
XX Claim 2: Page 116-151; 266pp; English.
XX The present invention relates to the cloning and sequencing of the gene
XX cluster encoding a modular type I polyketide synthase (PKS) enzyme
XX involved in the biosynthesis of the macrolide antibiotic nystatin.
XX The nystatin PKS is useful as antifungal antibiotics. The present
XX sequence is a Streptomyces noursei nys1 DNA of nystatin PKS gene cluster.
SQ Sequence 65140 BP; 8270 A; 25171 C; 22273 G; 9426 T; 0 other:
Query Match 5.6%; Score 99.2; DB 22; Length 65140;
Best local Similarity 43.5%; Pred. No. 1.8e-07;
Matches 724; Conservative 0; Mismatches 913; Indels 27; Gaps 5;
QY 62 ACCGTCGTCGCGGGGGGCGGACCTCATCGAGCGGACGACCCGGGGCGGCGG 121
DB 19698 ACGAGCGCTGCGCGGCGGACCTGCTTCAGTACGCGCCGCTTCAGGCGCGGCGG 19757
QY 122 ACCTGGCGTGGCGGCGGCGGACCGGATCGCGCGATCTGTGGAGCGCGCGCGC 181
DB 19758 CCTGGCGCGGCGGACCGAGGATCTGCGCGCGCGCGCGCGGCGGCGGACCG 19817
QY 182 ACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241
DB 19818 ACCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19877
QY 242 ACGAGCGACGACCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
DB 19878 ACGCGGACCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19937
QY 302 CGAGGCGTACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 361
DB 19938 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 19997
QY 362 CGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
DB 19998 TCACCATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20057
QY 422 ACCTGAGCGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481
DB 20058 CCGCGGAGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20117
QY 482 CAACGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
DB 20118 TCACAGTGAAGTGAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20177
QY 542 CGGCGATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
DB 20178 TCCTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20237
QY 602 GCGCCTTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 661
DB 20238 CCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 20297
QY 662 CGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 721

PS Claim 1: Page 188-254; 266pp; English.

XX The present invention relates to the cloning and sequencing of the gene
CC cluster encoding a modular type I polyketide synthase (PKS) enzyme
CC involved in the biosynthesis of the macrolide antibiotic nystatin.
CC The nystatin PKS is useful as antifungal antibiotics. The present
CC sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.

XX Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other:

Query Match 5.6%; Score 99.2; DB 22; Length 125401;
Best Local Similarity 43.5%; Pred. No. 1.7e-07;
Matches 724; Conservative 0; Mismatches 913; Indels 27; Gaps 5;

62 ACCGTGCTCTGCGGGGCGGACCTCATGACGGCAGACACCCCGGGGCGGGCGG 121
Db 79959 ACAGAGCCCTGCGCGCACTGCGCTTCCATGACGGCCCGCTTCCAGGGCGTGGCGG 80018
Oy 122 ACCTGGCGCTGCGGGGCGGACCGCATCGCGGCATGCTGTGGAGCGCGCGCGC 181
Db 80019 CCTGGCGCGCGGACACGAGGTCTACGCGGAGTGGCCCTGCGGACGGCGCGACCG 80078
Oy 182 ACACCGGGGTGACGCTGCGGGCTGTGGTGGCGCGCCCTTCAATGACTGCGACACCC 241
Db 80079 ACCCGCGCGCTTGGGACTGACCGCGCCCTGCTGAGCGCGGACAAACAGCGCGCGCT 80138
Oy 242 ACAGAGCACTACTCTGCTAGGCGTGGCGGACATGAGCGCCCAAGATCTGCGAGGGGTCA 301
Db 80139 ACAGGCACTGCGGGGCGCATACGCGCGCGCGCTGCGTGGCGGAGGGGTCTCGC 80198
Oy 302 CCAGCGGTGTACAGGGGCAATTGCGGCATACAGCTGCGCGCGCGTGGCGGACGCGCA 361
Db 80199 TCG 80258
Oy 362 CCG 421
Db 80259 TCACCATGCGCGCTTACGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 80318
Oy 422 ACCTGAGCGCGCTTGGGGGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 481
Db 80319 CCGCGGAGGTGCG 80378
Oy 482 CAACGCTGCG 541
Db 80379 TCCAGCTGCAATGAGACCGCGCTCCAGGGCGCGCGCGCGCGCGCGCGCGCG 80438
Oy 542 CGGCAATGCGGGGACCTGCGCGCGGAGAAAGCATGCGAGCGCGCGCGCGCGCGCG 601
Db 80439 TCTTCG 80498
Oy 602 GCGCGCTTACCG 661
Db 80499 CCG 80558
Oy 662 CGCTGAGCG 721
Db 80559 TCACCAACCTCTACACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 80618
Oy 722 TGGCGCGCGCGCGGAGAACTTCCGATCGCGCGCGCGCGCGCGCGCGCGCGCG 781
Db 80619 CCGCGCGCTGCG 80678
Oy 782 CGCACCAAGGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 841
Db 80679 TGGCTCTGTCACCG 80738
Oy 842 TCGAGGCG 901
Db 80739 GCTGATCTCGACCG 80798
Oy 902 CCACCATGCTCAACGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 961
Db 80799 CCGACCG 80852

Oy 962 AGCCCTTCCCGAAGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1021
Db 80853 ACAGAGCG 80912
Oy 1022 CCAAGTACGAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1081
Db 80913 TCCCGCTGCG 80972
Oy 1082 AACCGAGCG 1141
Db 80973 GCACCG 81029
Oy 1142 CGCAGACGAGCG 1201
Db 81030 GCAACCTGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81089
Oy 1202 ATGCG 1261
Db 81090 CCGAATCAGCG 81149
Oy 1262 CCG 1321
Db 81150 CCGC-----CGTGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81200
Oy 1322 TGTGCTGTGACCG 1381
Db 81201 TGTGCTGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81260
Oy 1382 GCG 1441
Db 81261 GCTGTGACACCGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81320
Oy 1442 TCACG 1501
Db 81321 GCACCTGCACTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81380
Oy 1502 CCTTAC-----ATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555
Db 81381 CGGACAGCG 81440
Oy 1556 CAACCGCTACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1615
Db 81441 CCGCGACCG 81500
Oy 1616 AGGAGTGTGCG 1672
Db 81501 TGACAGCAAGCG 81560
Oy 1673 AAGACCTGTGCG 1716
Db 81561 TGACCTTGGAACAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 81604

RESULT 7
AA073500/c
ID AA073500 standard; DNA: 8438 BP.

XX AA073500;

XX 15-MAY-1995 (first entry)

XX DNA encoding Pseudorabies virus large latency transcript.

XX Pseudorabies virus; PRV; LNV; large latency transcript;

XX attenuated virus; vaccine; early protein 0; EP0; HSV-1; ICP0;

XX protecting animals; deletion mutants; swine; ds.

XX Pseudorabies virus.

XX Key Location/Qualifiers

XX misc_feature 1..7013

XX FT misc_feature 1..7013

QY 559 GCGGAGGAAGCCATGGCCAGCGCGCCGATTCGACCGCGCCCTTACCGCCG 618
 Db 24029 ACCGTCGGCTGACGCTACACCGACCGCGCGCGCTGGCGCCCGGATCACCGAA 24088
 QY 619 GCGCGCCGCGCCACCAACGAGATCATGAGGTGTGCGCGCGCGTGAAGCGCATYGGC 678
 Db 24089 GCGCGCGCGCCACCAAGGCGCGCTGACAGCGGTGTGCTGCTGCTGCGCGCTGCGCCACCGCG 24148
 QY 679 GGCATCTACGCAACCAATGCGCGAGAGGCGAGACATCTGTGGCGCGCTGAAGAA 738
 Db 24149 GACGCGCGCGCCACCGCGCGCGCGCTGACCTGACCTACACCAACCGCGCTGACAGGCG 24208
 QY 739 ACCCTCCGATTCGCGCGCGCGAGCTGAGCTGCGGTGTGATCTGCAACCAAGGTGATG 798
 Db 24209 CTGCGCGCGCGCGCATGAGCGCGCGCTGTGAACTGACCGCGCGCGCGCTGCGCGCTG 24268
 QY 799 GCGCGCGCGCAATTCGCGCGCGCGCGAGACGCTGCGCGCTGATGCAAGCGCGCATGAGCG 858
 Db 24269 GCGCGCGCGCGCAAGGCTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 24319
 QY 859 GCGCGAGAGCTCTGCGCTGAGAGCGGTATCCTAGCTGCGCGCGCTGACCATGCTCAAGCAG 918
 Db 24320 GCGCGCGCGCGCTGCGCGCTGAGAACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 24379
 QY 919 GACCGCGCTGCTGCGCGCGCGCGAGCATGATGATGATGATGATGATGATGATGATGATG 978
 Db 24380 ACCCTGAGCG 24439
 QY 979 AGCGCGCGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1038
 Db 24440 GACCG 24496
 QY 1039 CCGCGAGCTGACG 1098
 Db 24497 GCG 24556
 QY 1099 ATGCTGCGCGCTGCG 1158
 Db 24557 GCG 24616
 QY 1159 CATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1215
 Db 24617 CACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24676
 QY 1216 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1275
 Db 24677 GACCTGCGAGAACTGCG 24736
 QY 1276 GCGCTGCG 1335
 Db 24737 GCGCTGCG 24796
 QY 1336 CCG 1392
 Db 24797 ACCG 24856
 QY 1393 ATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1452
 Db 24857 GTGCTGCG 24916
 QY 1453 CATGCG 1512
 Db 24917 CTGCGCGCGCTTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 24973
 QY 1513 GCGCTGAAGGCG 1563
 Db 24974 GGCACACTGACG 25024

RESULT 10
 AAL40781
 ID AAL40781 standard; DNA: 88421 BP.
 XX

AC AAL40781;
 XX 03-OCT-2002 (first entry)
 DT 88421nt genomic DNA of ramoplanin producing Actinoplanes sp.
 XX
 DE Ramoplanin: ramoplanin biosynthetic pathway polypeptide; antibiotic;
 KW biosynthesis gene cluster; bioengineering; peptide synthetase module;
 KW adenylation domain; hydroxyphenylglycine; Hpg; antibiotic precursor;
 KW chlorinate; lipopeptide; gene; ds.
 XX
 OS Actinoplanes sp.
 FH Location/Qualifiers
 FT 2077..3078
 CDS /tag- a
 FT /product= "Protein of ORF 1"
 FT 318..4032
 CDS /tag- b
 FT /product= "Protein of ORF 2"
 FT 4038..5048
 CDS /tag- c
 FT /product= "Protein of ORF 3"
 FT complement (6665..5814)
 CDS /tag- d
 FT /product= "Protein of ORF 4"
 FT complement (7703..6693)
 CDS /tag- e
 FT /product= "Protein of ORF 5"
 FT complement (9464..8130)
 CDS /tag- f
 FT /product= "Protein of ORF 6"
 FT 9691..10761
 CDS /tag- g
 FT /product= "Protein of ORF 7"
 FT complement (12751..10829)
 CDS /tag- h
 FT /product= "Protein of ORF 8"
 FT complement (13617..12802)
 CDS /tag- i
 FT /product= "Protein of ORF 9"
 FT complement (15203..13614)
 CDS /tag- j
 FT /product= "Protein of ORF 10"
 FT 15591..15863
 CDS /tag- k
 FT /product= "Protein of ORF 11"
 FT 15880..19035
 CDS /tag- l
 FT /product= "Protein of ORF 12"
 FT 19032..39713
 CDS /tag- m
 FT /product= "Protein of ORF 13"
 FT 39713..65800
 CDS /tag- n
 FT /product= "Protein of ORF 14"
 FT 65826..66530
 CDS /tag- o
 FT /product= "Protein of ORF 15"
 FT 66546..67370
 CDS /tag- p
 FT /product= "Protein of ORF 16"
 FT 67384..70059
 CDS /tag- q
 FT /product= "Protein of ORF 17"
 FT 70099..70662
 CDS /tag- r
 FT /product= "Protein of ORF 18"
 FT 70659..71906
 CDS /tag- s
 FT /product= "Protein of ORF 19"
 FT complement (73439..71964)
 CDS /tag- t


```
Db 46830 TCATGACCTGCGAGCGCCGCTGCTGCTGCTGCGGCTGCGGCGGCGGCGACCGGACG 46889
Qy 689 CCACCCATGTCGCGAGGAGGACATCTGTCGCGGCTGAGGAAACCTTCCGCA 748
Db 46890 TCCGATCGGAGCGGTGTCGCGCGCGCGCGGCGGCGGTGAGAACCTGTCGCT 46949
Qy 749 TCGGCGCGGAGCTGAGCTGCGGCTGATCTGCGACACACAGGTATGAGCCAGCCCA 808
Db 46950 TCTTCGTCAACACCTCTGCTGCGCACCGACCTGACCGGCGGCGGCGGCTGACCGACG 47009
Qy 809 ATTTCGGCGGCTGCGGAGAGCGTGCCTGATGAGAGCGCGGCGGCGGCGGCGGAGAGC 868
Db 47010 TGTCTGCGGAGCTGCGGAGCTGACCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCTG 47069
Qy 869 TCTGCTGAGACGCTTACCTTACCTGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCTG 928
Db 47070 AGAAGCTGCTGAGAGAGCTACCCCGCGGCTC-----GCTGCGCGGCGGCGGCGGCTG 47123
Qy 929 TGTGCGCGGAGCGACCATCATCTGCTGTCAGAGCCCTTCCCGGAACTGAGCGGCGCG 988
Db 47124 TCCAGGTCTATGCTGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 47183
Qy 989 ACCGAGTGAAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1048
Db 47184 CGATGCTGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 47243
Qy 1049 AGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 1108
Db 47244 CCGAGACCTTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 47303
Qy 1109 TCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1168
Db 47304 CCGACCTTTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 47363
Qy 1169 TGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1228
Db 47364 GCGTCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 47411
Qy 1229 TGGAGCGGCGGCTATGAGATGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1288
Db 47412 TGGAGCGGAGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 47471
Qy 1289 GCGGCGGAGCTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1348
Db 47472 CCGAGATCACGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 47531
Qy 1349 CCGATACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1408
Db 47532 CCGTACCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 47591
Qy 1409 TCAACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1432
Db 47592 TCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 47615

RESULT 11
AAS08693
ID AAS08693 standard; DNA; 109519 BP.
XX
XX AAS08693;
DT 26-SEP-2001 (first entry)
XX
XX Micromonospora DNA encoding biosynthetic enzymes for Evernimycin.
DE Evernimicin; antibiotic; bottle-neck gene; orthomycin.
KW fermentation; ds.
XX
XX Micromonospora carbonacea var. africana.
OS
XX
XX Key Location/Qualifiers
FH complement (132..1382)
FT CDS
FT /*tag= a
```

```

FT      CDS      /*tag- ab
FT      CDS      22748..24172
FT      CDS      /*tag- ac
FT      CDS      /*product- "Evrg"
FT      CDS      22736..22740
FT      CDS      /*tag- ad
FT      CDS      complement (24177..25223)
FT      CDS      /*tag- ae
FT      CDS      /*product- "Evrv"
FT      CDS      complement (25230..25233)
FT      CDS      /*tag- af
FT      CDS      25550..26626
FT      CDS      /*tag- ag
FT      CDS      /*product- "Evri"
FT      CDS      26685..30479
FT      CDS      /*tag- ah
FT      CDS      /*product- "Evri"
FT      CDS      26672..26676
FT      CDS      /*tag- ai
FT      CDS      complement (30557..31876)
FT      CDS      /*tag- aj
FT      CDS      /*product- "Evri"
FT      CDS      complement (31885..31888)
FT      CDS      /*tag- ak
FT      CDS      complement (31941..32882)
FT      CDS      /*tag- al
FT      CDS      /*product- "Evri"
FT      CDS      complement (33167..34405)
FT      CDS      /*tag- am
FT      CDS      /*product- "Evrm"
FT      CDS      complement (34414..34418)
FT      CDS      /*tag- an
FT      CDS      complement (34449..35210)
FT      CDS      /*tag- ao
FT      CDS      /*product- "Evri"
FT      CDS      complement (35219..35221)
FT      CDS      /*tag- ap
FT      CDS      complement (35294..36238)
FT      CDS      /*tag- aq
FT      CDS      /*product- "Evri"
FT      CDS      complement (36235..36963)
FT      CDS      /*tag- ar
FT      CDS      /*product- "Evri"
FT      CDS      complement (36998..38026)
FT      CDS      /*tag- as
FT      CDS      /*product- "Evri"
FT      CDS      complement (38072..38566)
FT      CDS      /*tag- at
FT      CDS      /*product- "Evri"
FT      CDS      complement (38892..40163)
FT      CDS      /*tag- au
FT      CDS      /*product- "Evri"
FT      CDS      complement (40216..40890)
FT      CDS      /*tag- av
FT      CDS      /*product- "Evri"
FT      CDS      complement (40899..40902)
FT      CDS      /*tag- aw
FT      CDS      complement (40887..41576)
FT      CDS      /*tag- ax
FT      CDS      /*product- "Evri"
FT      CDS      complement (41679..42707)
FT      CDS      /*tag- ay
FT      CDS      /*product- "Evri"
FT      CDS      complement (42714..42717)
FT      CDS      /*tag- az
FT      CDS      complement (42810..43799)
FT      CDS      /*tag- ba
FT      CDS      /*product- "Evri"
FT      CDS      complement (43807..43811)
FT      CDS      /*tag- bb
FT      CDS      complement (43799..44866)
FT      CDS      /*tag- bc
FT      CDS      /*product- "Evri"

```

```

FT      CDS      complement (45014..45760)
FT      CDS      /*tag- bd
FT      CDS      /*product- "Evri"
FT      CDS      complement (45767..45770)
FT      CDS      /*tag- be
FT      CDS      complement (45962..46714)
FT      CDS      /*tag- bf
FT      CDS      /*product- "Evri"
FT      CDS      complement (45952..45956)
FT      CDS      /*tag- bg
FT      CDS      complement (47156..49234)
FT      CDS      /*tag- bh
FT      CDS      /*product- "Evri"
FT      CDS      51627..52715
FT      CDS      /*tag- bi
FT      CDS      /*product- "Evri"
FT      CDS      51629..51622
FT      CDS      /*tag- bj
FT      CDS      52889..53557
FT      CDS      /*tag- bk
FT      CDS      /*product- "Evri"
FT      CDS      53554..54207
FT      CDS      /*tag- bl
FT      CDS      /*product- "Evri"
FT      CDS      complement (54362..55117)
FT      CDS      /*tag- bm
FT      CDS      /*product- "Evri"
FT      CDS      complement (55125..55128)
FT      CDS      /*tag- bn
FT      CDS      /*product- "Evri"
FT      CDS      complement (55135..56094)
FT      CDS      /*tag- bo
FT      CDS      /*product- "Evri"
FT      CDS      complement (56100..56103)
FT      CDS      /*tag- bp
FT      CDS      complement (56184..56813)
FT      CDS      /*tag- bq
FT      CDS      /*product- "Evri"
FT      CDS      56961..58709

```

Query Match 4.8%; Score 84.6; DB 22; Length 109519;
 Best Local Similarity 45.1%; Pred. No. 3.2e-05;
 Matches 542; Conservative 0; Mismatches 629; Indels 30; Gaps 5;

```

QY 117 CGCGACCTGGGCGTGGCGGCGACCGCATGCGCGCATGTCGAGCGCG 176
D 94498 CGTGACCTGTCACGCGCGCGCGCGCATGCGCGCATGTCGAGCGCG 94557
QY 177 CGCGACACCGCGGCGTGCAGCTGTCGCGCGCGCGCGCGCGCGCGCG 236
D 94558 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94617
QY 237 CACCGACGACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 236
D 94618 CGGCTTGGCGCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94677
QY 297 CGTCACACGCGTGTGTCACGCGCATTTGCGCGCATGCGCGCGCGCG 356
D 94678 GGTACCGCGCGTGTGTCACGCGCGCGCGCGCGCGCGCGCGCGCGCG 94737
QY 357 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 404
D 94738 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94797
QY 405 CGACCGCTTGGCGCGTGTGTCACGCGCGCGCGCGCGCGCGCGCGCG 464
D 94798 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94857
QY 465 CTGTATGCTGGCGCGCATTCAGCGCGCGCGCGCGCGCGCGCGCGCG 524
D 94858 CGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 94917
QY 525 CACCGACGAGGAAATCGCGCGCATGCGCGCGCGCGCGCGCGCGCGCG 584

```

D	94918	CTTGACGCGGTCAACCGTGGGATACC66GCCAGCACGCCCCGGCCTTGACCGGTTAC	94977
OY	585	CATCGGGCATTTTCGACCCGGGCCTTCTACTACCGCCGCCGCCGCCACCAACGAAGAT	644
D	94978	CTTGAGACTGCCCGGCGCGCCGGGTGCCGTGGTGGGGCGAAGCGGCCGCCGGAGAAG	95037
OY	645	CATGAGGTGTGCGGGCCGTGAGCGCGCATGGCGCATTTACGCCAACCCACATGCGGA	704
D	95038	CACCTCGCCCGCTCTCTCACCGGCACAGTGCAGACCCGAGAGGGCCGGTCTACCTTGA	95097
OY	705	CGAAGCGAGACAATCTGTGGCCGCGCTGGAGAAACCTTCCGATCGGCCCGAGCTGGA	764
D	95098	CGGGGCGCACTGTGCGGCTACCCGCTGAGAGAACTGCCCCGGGCCGTGCGGCTGCT	95157
OY	765	CGTGGCGGTGTGATCTTCGACACACAGATGATGGGCCAGCCATTTCGGCG- ----	818
D	95158	CGCGGAGCGTACGCTTTCTTCACGCGCACAGTCCGGGAGAACTGTGCTGCGCGCCGC	95217
OY	819	CTMGCGGAGACGCTGGCCGCTGATGAGAGGCCACCATGGCCGGCCAGAGAGTCTGCTGA	878
D	95218	CGCCGACGAGGGGAGCTGACCGCGGCACACCGGGCGGCCGCGCTGTGATGGGTGCA	95277
OY	879	GCGGTATCCCTACGTGGCGCGGCTCCACCATGCTCAAGCA- ---GGACCGCGTGTGTGCG	934
D	95278	CGCCACGCGCGGGGTGGGACACCGTGTGGGCGAGGAGGGCGGACAGCTCTCCGGCGG	95337
OY	935	CCGACGCAACCATATCATCTGTGTGCAAGCCT-- TCCCCGAACTGAGCGGGCGGACCT	992
D	95338	CGAGCGGAGGCGCTTCGCGCTGGGCCCGGCGCTGCTCTCCCGCGCCGGGATGCTGTGCT	95397
OY	993	GGATGAATCTCGCGCGCGAGCGCGGCAAAATCAAGTAGCAGCGTGGTC-----CCGAGCT	1046
D	95398	CGACGACCGGACGAGGGGGCTGACCCGCTCCGCCCGCAGCGGGGTGCTGCTCGGCCCT	95457
OY	1047	GCAAGCGGCGCGCCCATCTACTTATGATGAGACGAAACCCGAGTGCAGCGCATCTTGGC	1106
D	95458	GGCGCGACCGCCCGCGGACATCTCGCTGTGATACGCCACCGGCTCAGGGGCTCGC	95517
OY	1107	GTTCCGCCGACCATGATCGGCTCGGAGGGCTGCGGACGACGAGAGCGCCGATCGCG	1166
D	95518	CGACCTCGACGAGATCTGTGTCTCTGACGCGCGCGGGGTGTCAGCGTGGCGCGGACGA	95577
OY	1167	CCTGTGGGGACCTTCCGCGCGGCTCTGGGGCATATGCGCGCGACATGCGCTTCTCC	1226
D	95578	CGAGTTGGTCGCGCGCGCGGTGTGTACCGGGACGAGTGGCTGCTCAAGGAGCGGCCGA	95637
OY	1227	GCTGAGAGAGCGGCGTATGGAAATGACCGGCGCTGACCGCGCGCGGCTTGGGCTGAGCCG	1286
D	95638	GGCGGGGTAACCTGCGCCCTGACGCGCCGCCCTGAGCGCGGCTTCGGGATTTCCCCGAGCGC	95697
OY	1287	G	1287
D	95698	G	95698
RESULT 12			
AAZ87286	ID	AAZ87286 standard; DNA; 1248 BP.	
XX	AC	AAZ87286;	
XX	DT	05-JUN-2000 (first entry)	
DE	XX	S. venezuelae desosamine biosynthetic gene desi, SEQ ID NO:7.	
KX	XX	Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;	
KM	XX	neomethymycin; narbomycin; polylhydroxalkanate monomer synthesis;	
KW	XX	biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;	
KV	XX	chronic obstructive pulmonary disease; respiratory inflammation;	
KW	XX	hypercholesterolaemia; crop protection agent; ds.	
OS	XX	Streptomyces venezuelae ATCC15439.	
XX	XX		

FH	Key	Location/Qualifiers
FT	CDS	1..1248
FT		/*tag= "A
XX		/product= "DesI"
FM		
PM	WO200000620-A2.	
XX		
PD	06-JAN-2000.	
XX		
PF	25-JUN-1999; 99WO-US14398.	
XX		
PR	26-JUN-1998; 98US-0105537.	
XX		
PA	(MINU) UNIV MINNESOTA.	
PI	Sherman DH, Liu H, Xue Y, Zhao L;	
DR	WPI: 2000-160679/14.	
XX	P-PsDB: AAY77181.	
PT	Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.	
PS	synthesis of methymycin and pikromycin -	
XX		
XX	Claim 3; Page 353-354; 438pp; English.	
CC	The invention relates to an isolated and purified nucleic acid segment comprising a desosamine biosynthetic gene cluster, a fragment or its biologically active variant, where the nucleic acid sequence is not derived from the eryC gene cluster of <i>Saccharopolyspora erythraea</i> or streptomycetes antibioticus. The invention also relates to a macrolide biosynthetic gene cluster, or fragments thereof. The macrolide biosynthetic gene cluster encodes proteins which synthesise methymycin, pikromycin, neomethymycin, narbomycin or a combination of these compounds. Recombinant or augmented cells comprising the desosamine and/or macrolide biosynthetic gene clusters are useful for the production of biologically active macrolides. The macrolide biosynthetic proteins are useful for synthesis of methymycin, pikromycin, neomethymycin and narbomycin. The alternative termination of polyketide synthase may be useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA) monomers. The compounds produced by the recombinant host cells are useful as biopolymers, e.g., in packaging or biomedical applications, to engineer PHA monomer synthases or to prepare biologically active agents, such as chemotherapeutics,	
CC	immunosuppressants, agents to treat asthma, chronic obstructive pulmonary disease as well as other diseases involving respiratory inflammation, cholesterol-lowering agents or macrolide-based antibiotics which are active against a variety of organisms, e.g., bacteria, including multi-drug resistant pneumococci and other respiratory pathogens, as well as viral parasitic pathogens, or as crop protection agents (e.g., fungicides or insecticides) via expression of polyketides in plants.	
CC	Sequences AA287286-287294 represent desosamine biosynthetic genes from <i>Streptomyces venezuelae</i> ATCC 15439, which encode proteins AAY77181-vY7189.	
CC		
XX		
SO	Sequence 1248 BP; 177 A; 546 C; 377 G; 148 T; 0 other:	
Query Match	4.7%; Score 81.8; DB 21; Length 1248;	
Best Local Similarity	44.2%; Pred. No. 0.00015;	
Matches 439; Conservative	0; Mismatches 542; Indels 12; Gaps 2;	
OY	32 CATATGCCCAATCCGATTCCCGACGGCTTGACTGTCTGCAGGGGCGCACCCCTCATCG 91	
DB	215 CCACCTGCACAACGCGACGCGCGGGCTCCAGCTCTTCGCGCAGCGCGCGGCGCTTACCGCGG 274	
OY	92 ACGGCAACAACACCCCGGGGCGGCGCGCGACTGTGGCGCTGCGGCGGACCGCATCGCGC 151	
DB	275 AAGTGAATCATGCGCTGATGATGATGCTTGGCCGACCCCGACAGCACTGCGCTGGATCGGGC 334	
OY	152 CCAATGGCGCATCTGTGTGGAGACCGCGCGCGACAAACCGGGTGCACGTGTGGGGCTGTGTC 211	
DB	335 TCACCCCGGCTTTCGCGCATTCGACCCGGAACCGGCAACCTTCGACCCGCGACCAAGTTGG 394	
OY	212 TCGCGCGCGGCTTCATTCGACTGCACACCAGACGACGACCAACTACTGTGCTCAGCGCTCGCG 271	

```

Db 395 CCGCCGCGGTACACCCCGACCTCGCGCGGTGCGGCTCCACCTCTGGGGCGCCCT 454
Oy 272 ACATAGCGCCCAAGATCTCGACAGGGGTGTCACGAGGTGTACAGGGCAATTCGGCATCA 331
Db 455 GCGCCGCGGACGAGCTGCGGAAGTCCGCGACGACGCGCTGCGGCTGTACTTGCAGC 514
Oy 332 GCGTGGCGCGCTGGCGACGCAACCGCGCCCGCTGGACCTGCGTGGAGAGGG 391
Db 515 CCGCGGACGCGCTCGCGCTCGCGGTGAGAGCGCGCGCGCGCGCGCGCGCGCG 574
Oy 392 GCTCTTACGCTTTCGAGCGCTTTCGCGACTTACCTGAGCGCTTGGCGGCGCGCG 451
Db 575 AGGTCTTACGCTTTCGAGCGCTTTCGCGACTTACCTGAGCGCTTGGCGGCGCG 631
Oy 452 CCGTAAAGCGCGCTGTATGTTGGGCGCATTCAGCTGCGCGCGCGCGCGCGCG 511
Db 632 TCACGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 691
Oy 512 TCACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 571
Db 692 TCACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 751
Oy 572 TGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 631
Db 752 TGGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 811
Oy 632 CCACGAGAGAGATCATGAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 691
Db 812 CCGCGTACCGCGAGCAGCTCGCGGAGCTCGCGCGCGCGCGCGCGCGCGCG 871
Oy 692 CCGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 751
Db 872 ACAGCGCTCAACACGAGTACGATGCTGAGATGAGACGCGCGCGCGCGCGCG 931
Oy 752 GCGCGAGCTGAGCTGCGCGGTGTGATCTGCGACACAGGTGATGCGCGCGCG 811
Db 932 ACCGCGAGCTGCTCATGAGAGGTCTGAGAGCGCGAGCGGTGACACCGCGCG 991
Oy 812 TCGCGCGCTGCGCGAGAGCGGTGCGCGGTGATGAGCGCGCGCGCGCGCGCG 862
Db 992 CCGCGGCGCTGCGCGAGAGCTGAGAGCGGTGCGCGCGCGCGCGCGCGCGCG 1051
Oy 863 AGAGCGTCTGCTGAGACGCTATCCCTAGTGGCGCGCGCGCGCGCGCGCGCG 922
Db 1052 CCGAAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111
Oy 923 GCGTGTCTGCTGCGCGAGCGACATCATCATCTGCTGAGCGCGCGCGCGCGCG 982
Db 1112 ACATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1171
Oy 983 GCGCGAGCTGATGAGTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1015
Db 1172 CCGCGCAGCGCGAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1204

```

```

RESULT 13
AAD39044
ID AAD39044 standard; DNA: 1248 BP.
AC AAD39044:
XX 23-SEP-2002 (first entry)
XX
DE Streptomyces venezuelae Desi gene.
XX
KW Glycosylated polyketide; modified recombinant bacterial host cell;
KW mRbHC; macrolide; anthracycline; angucycline; avermectin; milbemycin;
KW tetracycline; polyene; polyether; ansamycin; isochromanone; sugar;
KW desosamine; Desi; 4-dehydrase; enzyme; gene; ds.
XX
OS Streptomyces venezuelae.
XX

```

```

FH Key Location/Qualifiers
FT CDS 1..1248
FT /tag- a
FT /product- "Streptomyces venezuelae Desi protein"
FT /trans- except- (pos:1..3, aa:net)
FT /note- "CDS does not include start codon"
FT /partial
PN WO200229035-A2.
PD 11-APR-2002.
XX 05-OCT-2001; 2001WO-US31255.
PF 05-OCT-2000; 2000US-238185P.
PR (MINU) UNIV MINNESOTA.
PA (LIUH/) LIU H.
PA (SHER/) SHERMAN D H.
PA (ZHAO/) ZHAO L.
PI Liu H, Sherman DH, Zhao L;
XX WPI: 2002-405171/43.
XX P-PSDB: AAE24229.
XX
PT Modified recombinant bacterial host cells in which the expression and
PT activity of nucleic acids encoding sugar biosynthetic enzymes has been
PT altered, useful for producing metabolites with altered sugar structures
PT
PS Disclosure: Page 153-154; 174pp; English.
XX
CC The invention provides a method to alter the sugar structure diversity
CC for a particular metabolite via the recruitment and collaborative action
CC of sugar genes from a variety of sugar biosynthetic pathways to yield a
CC metabolite comprising a non-natural sugar, e.g., a novel glycosylated
CC polyketide. The invention also relates to a modified recombinant
CC bacterial host cell (mRbHC) in which the expression and activity of
CC nucleic acids encoding sugar biosynthetic enzymes has been altered.
CC The mRbHCs may be cultured to produce the modified sugar products.
CC e.g. a macrolide, anthracycline, angucycline, avermectin, milbemycin,
CC tetracycline, polyene, polyether, ansamycin or isochromanone.
CC The present sequence is Streptomyces venezuelae sugar (desosamine)
CC biosynthetic gene cluster Desi (4-dehydrase) gene.
XX
S0 Sequence 1248 BP; 177 A; 546 C; 377 G; 148 T; 0 other;

```

```

Query Match 4.78; Score 81.8; DB 24; Length 1248;
Best Local Similarity 44.28; Pred. No. 0.00015;
Matches 439; Conservative 0; Mismatches 542; Indels 12; Gaps 2;

```

```

Oy 32 CCATGTCCCAATCCGATTCGCGCGCGCTTCGACGTGCTGCGCGGGGCGACCGCTCATCG 91
Db 215 CCACCTGCAAGCGCACGCGCGCGCTTCAGCTCTTCGCGACGCGCGCGCGCTTCACGGCG 274
Oy 92 ACGGAGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 151
Db 275 AAGTATCATGCTCGTGCATGATGACGTTGCGCGCGCGCGCGCGCGCGCGCGCG 334
Oy 152 CCATGCGGATCTGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 211
Db 335 TCACCGCGGCTTTCGCGCATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 394
Oy 212 TCGCGCGCGCTTCATGCTGCGACACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 271
Db 395 CCGCGCGCGTACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 454
Oy 272 ACATGAGCGCGCAAGATCTCGAGAGCGCTTCACACGAGTGTGACGCGCGCAATTCGGCGCATCA 331
Db 455 GCGCGCGCGACGACGCTGCGGAAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 514
Oy 332 GCGTGGCGCGCGCTGCGCGACGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 391

```

Db	515	CCGCGACGCCCTCTCGGCTTCGGGTGCAGAGGCGCCGCCGCCGCGAGCCTGGGAGACGCCG	574
Qy	392	GCTCTTACCGTTTCGAGCGCTTCCGCGACTACTTGAGCGGTTGCGGCCACGCGCGCG	451
Db	575	AGGTTCTTACGCTTCCA---CGCCACAAGAGCGCTTCAACGCTTGTGAGGAGGCGCGCTG	631
Qy	452	CGGTCAAGCGCGGTGTATGTGGGGCCATTCAGACGTTCGCGCGCGCGGTATGCGGACT	511
Db	632	TCACCGGACGCGCCGACCTTCGGCGCCGGATTCGCGCCTTCAACAATTGGCTTGACC	691
Qy	512	TGCAGCGCGCGCCGACGACGAGAAATCGCGCCATCGCGGACCTTGCGGAGGAACCA	571
Db	692	TGCGCGGCGGACGCGCGCGCGCGGAGAACGCAACGCAATGAGGAGAGCGCGCGCCA	751
Qy	572	TGCGCAAGCGCGCCATTCGCAATTTGACCGCGCGCTTCTATCCGCGCGCGCGCGCA	631
Db	752	TGGCGCTTCACTCCCTCCGAGCGGTTTCCGAGGTATGACGACCGGACCGCGCAACGAC	811
Qy	632	CCACCGAAGATCATCGAGTGTGTCGCGCGCTGAGCGCGCATGCGGCACTTACGCCA	691
Db	812	CCGCGTACCGCGAGACACTTCGGGAGCTCCCGCGGTCTGTGCGCGACGACGACGCC	871
Qy	692	CCCAACATGCGCGACGAGGCGAGCAACATCTGCGCGCTGAGGAAACCTTCGCACTG	751
Db	872	ACGCGCTTCAACCAACCAACAGTACGTGATGTCGAGATGACGAGGCGCAACCGGCTAC	931
Qy	752	GCGCGAGCTGAGGAGCGCGGTGTGATCTCGCACACAGGTACATGGGCGGACGCCAAT	811
Db	932	ACGCGGCTTCTGATGAGAGTCTGAAAGCCGGAAGCGCTGACACCCGCGCTACTTCT	991
Qy	812	TCGCGCCCTGCGCGAGACGCTGCGCTGATGAG-----GCGGCATGCGCGCGC	862
Db	992	CGCGGCGCTCCACACGAGCTGAGACCGCTTACCGGCGACGCGACGCGCGCGCGACA	1051
Qy	863	AGGAGCTCTGGCTGAGAGCGGTATCCCTACGTGCGCGCTCCACCATGCTCAAGAGGAC	922
Db	1052	CCGAGCGCTTCGCGCGCGCGCGGTCTCTCTCCGACGCGACCGCACTTCGGCGAGCG	1111
Qy	923	GCGTCTGCTGCGCGCGGACGACCACTCATCCTGTGTGCAAGCCCTTCCCGAAGTGAAG	982
Db	1112	ACATCGCGCGGCTGCGGACCTGCTGCTGCGCGACCGCGCGCGCGGCGGACGCTGACG	1171
Qy	983	GCGCGACCTGTGATGAAATGCGCGCGCGGACCGG	1015
Db	1172	CGCGCACCGCGACGCGCGCGCGCGCGCTG	1204
RESULT 14			
AAZ87284/C			
ID	AAZ87284, standard; DNA: 12441 BP.		
XX			
AC	AAZ87284;		
XX			
DT	05-JUN-2000 (first entry)		
XX			
DE	S. venezuelae desosamine biosynthetic gene cluster pikB, SEQ ID NO:3.		
XX			
KW	Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin		
KW	nemeohemycin; narbomycin; polyhydroxyalkanone monomer synthase;		
KW	biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;		
KW	chronic obstructive pulmonary disease; respiratory inflammation;		
KW	hypercholesterolaemia; crop protection agent; ds.		
XX			
OS	Streptomyces venezuelae ATCC15439.		
XX			
PN	WO200000620-A2.		
XX			
PD	06-JAN-2000.		
XX			
PF	25-JUN-1999; 99MO-US14398.		
XX			
PR	26-JUN-1998; 98US-0105537.		

[illegible]

Oy	432	CCGTCACAGCGCGCTGTATGGTGGGCCATTCAACGCTGGCGGCGCGGCTCATGCGGACT	511
Db	11565	TCAACGAGCGCCACCACCTCAGCGCGCGGATCCGCGCCCTCCACAATCTTGCGCTTCAGGC	11506b
Oy	512	TGCACGCGCGCGCCACCGCACGAGAAATTCGGCGCCATCGCGGACCTGGCCGAGGAAGCA	571
Db	11505	TGCCGCGCGGCGACGCCCCCGCGCGGAGCCACAGCCAATGTGCGAAGCGCGCGCGGCA	11446b
Oy	572	TGGCAGCGCGCGCCATTCGGCATTTTGGACCGGCGCTTCACCGCGCGCGCGCGCGGCA	631
Db	11445	TGGGCTCTACCTCCCTCGACGGGTTTTCCGAGAGTATGACCGGANAACGGCGCCACACAG	11386b
Oy	632	CCACGGAAGAATCATTCGAGGTGTCCGGCGCGTGAAGCGCGCATGGCGCATCTACGCCA	691
Db	11385	CGGCTTACCGCGAGCACCTCGGCGGACCTCCCGGCGTCTCTCGCCGACACGACGCGC	11326b
Oy	692	CCACATGCGGCGAGGAAGCGGACATCTGTCGGCGCGCTGAGGAACCTTCGCGCATCG	751
Db	11325	ACGGCTCTAACACACCAACCAAGTACGTGATGTGAGATGACAGGACACACACCGGATCTC	11266b
Oy	752	GCGCGAGCTGGACCTGCGCGGTGTGTATCTGCACACAAAGTCTATGGGCCAGCCCAATT	811
Db	11265	ACCGGACCTGTCTATGTAGAGGTCTGAAAGCGCGAAGGGGTGACACCGCGCGCTACTCT	11206b
Oy	812	TCGCGCGCTCGCGCGAGACGCTGCCGCGTAGTCAG-----GCCGCATAGCGCGCGC	862
Db	11205	CGCCGCGGTGCGCACACAGCTGAGACCGCTACCGCGGCGACGCGCACCCCGCTGCGCGACA	11146b
Oy	863	AGGAGCTCTCGCTGGACGCGTATCCCTACGTGGCGGCGTCCACCATGTGTCACAGCGAGNC	922
Db	11145	CCGAACGCTTCGCGCGCGCGCGCTGTCTCTCTCGCACCGGACACCGCCATGGCGAGACG	11086b
Oy	923	GCGTCTGCTGGCGCGAGACGACACATCATCTGGTGTGCAAGCTTCCTCCGCAACTGAGCG	982
Db	11085	ACATCCGCGCGGTGCGCGCGACGTCGTGCGTCTGTGGCGGACCGCGGCGCGNACTGACCG	11026b
Oy	983	GCGCGGACCTGATATAAGTCGGCGCGGACGCGG 1015	
Db	11025	GCGCGCCAGCGACAGCGGCGCCCGCGCGCTG 10993	

Result	ID	AA87319/c	standard; DNA; 13613 BP.
XX	AA87319;		
XX	05-JUN-2000	(first entry)	
DE	S. venezuelae	desosamine biosynthetic gene cluster pikB.	
XX	Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;		
KW	neomethylmycin; naphthomycin; polyhydroxyalkanoate monomer synthase;		
KW	biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma;		
KW	chronic obstructive pulmonary disease; respiratory inflammation;		
XX	hypercholesterolaemia; crop protection agent; ds.		
OS	Streptomyces venezuelae	ATCC15439.	
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..809	
FT		/*tag= a	
FT		/partial	
FT		/product= "pikB gene cluster protein #1 (AAV77204)"	
FT		/note= "No initiation codon given in the specification."	
FT	CDS	806..2014	
FT		/*tag= b	
FT		/product= "pikB gene cluster protein #2 (AAV80998)"	
FT		2162..6741	
FT		/*tag= c	
FT	CDS		
FT		/product= "pikB gene cluster protein #3 (AAV77205)"	
FT		6834..7402	
FT	CDS		

FT		/+tag- d	
FT		/product= "pkb gene cluster protein #4 (AAV77206)"	
FT		/transl_except= (pos:6837..6841, aa:Gln)	
FT	CDS	7492..8205	
FT		/+tag- e	
FT		/product= "pkb gene cluster protein #5 (AAV77207)"	
FT	CDS	complement (7942..8205)	
FT		/partial	
FT		/+tag- f	
FT		/product= "pkb gene cluster protein #6 (AAV77208)"	
FT		/note= "no termination codon given in the specification"	
FT		/transl_except= (pos:8270..8272, aa:Val)	
FT		/transl_except= (pos:8273..8275, aa:Phe)	
FT		/transl_except= (pos:8276..8278, aa:Gly)	
FT	CDS	complement (10126..11139)	
FT		/+tag- g	
FT		/product= "pkb gene cluster protein #7 (AAV80999)"	
FT		complement (11271..12149)	
FT	CDS	/+tag- h	
FT		/product= "pkb gene cluster protein #8 (AAV77209)"	
FT	CDS	complement (12342..13799)	
FT		/+tag- i	
FT		/product= "pkb gene cluster protein #9 (AAV77210)"	
FT		complement (13706..15043)	
FT	CDS	/+tag- j	
FT		/product= "pkb gene cluster protein #10 (AAV77211)"	
FT		15404..15574	
FT	CDS	/+tag- k	
FT		/product= "pkb gene cluster protein #11 (AAV77212)"	
FT			
PN		WO200000620-A2.	
PD		06-JAN-2000.	
XX			
XX		25-JUN-1999; 99MO-USL4398.	
XX			
PR		26-JUN-1998; 98US-0105537.	
PA	(MINU) UNIV MINNESOTA.		
PI	Sherman DH, Liu H, Xue Y, Zhao L;		
XX	Wpi: 2000-160679/14.		
DR	P-PSDB: AAV77204, AAV77205, AAV77206, AAV77207, AAV77208,		
XX	AAV77209, AAV77210, AAV77211, AAV77212, AAY80998, AAY80999.		
PT	Desosamine and macrolide biosynthetic gene clusters, useful for, e.g.,		
PS	synthesis of methymycin and pikomycin -		
XX	Disclosure: Figure 32; 438pp: English.		
CC			
CC	The invention relates to an isolated and purified nucleic acid segment		
CC	comprising a desosamine biosynthetic gene cluster, a fragment or its		
CC	biologically active variant, where the nucleic acid sequence is not		
CC	derived from the eyfC gene cluster of Saccharopolyspora erythraea or		
CC	Streptomyces antibioticus. The invention also relates to a macrolide		
CC	biosynthetic gene cluster, or fragments thereof. The macrolide		
CC	biosynthetic gene cluster encodes proteins which synthesize methymycin,		
CC	pikomycin, neomethymycin, narbomycin or a combination of these		
CC	compounds. Recombinant or augmented cells comprising the desosamine		
CC	and/or macrolide biosynthetic gene clusters are useful for the production		
CC	of biologically active macrolides. The macrolide biosynthetic proteins		
CC	are useful for synthesis of methymycin, pikomycin, neomethymycin and		
CC	narbomycin. The alternative termination of polyketide synthesis may be		
CC	useful to prepare novel antibiotics and polyhydroxyalkanoate (PHA)		
CC	monomers. The compounds produced by the recombinant host cells are		
CC	as biopolymers, e.g., in packaging or biomedical applications, to		
CC	engineer PHA monomer synthases or to prepare biologically active agents,		
CC	such as chemotherapeutics, immunosuppressants, agents to treat asthma,		
CC	chronic obstructive pulmonary disease as well as other diseases involving		
CC	respiratory inflammation, cholesterol-lowering agents or macrolide-based		
CC	antibiotics which are active against a variety of organisms, e.g.,		
CC	bacteria, including multi-drug resistant pneumococci and other		

CC respiratory pathogens, as well as viral parasitic pathogens, or as crop
CC protection agents (e.g., fungicides or insecticides) via expression of
CC polyketides in plants. The present sequence represents the desomine
CC biosynthetic gene cluster (pik) from *Streptomyces venezuelae* ATCC
CC 15439, as given in figure 32.

XX Sequence 13613 BP, 1858 A; 4733 C; 5092 G; 1930 T; 0 other:

Query Match 4.7%; Score 81.8; DB 21; Length 13613;

Best Local Similarity 44.2%; Pred. No. 0.00011;

Matches 439; Conservative 0; Mismatches 542; Indels 12; Gaps 2;

QY 32 CCATGTCCTCCATTCGATTCGACCTTCGACCTGCTGTCGCGGGGCGACCTTCATCG 91
DB 12669 CCACCTGCAACGCGACGCGGGCTCCAGCTCTCGCGGACGCGCGGCTCCACCGGCG 12610
QY 92 AGGCGACCAACACCCCGGGGCGCGCCCACTGGGCGCTGCGGCGACCGCATCGCG 151
DB 12609 AAGTGATCATCGCGTGATGACCTTGCGCCGACCCCGACGCACTGCGCTGATCGGCC 12550
QY 152 CCATCGGCGATCTGTGAGCGCGCGCGACACCCGCGGTGAGCTGTGCGGCTGTGG 211
DB 12549 TCACCCCGGCTTTCGCGCATGATGACCGGACACCGGCAACCTGACCCGACCGTGG 12490
QY 212 TCGCGCCCGGCTTCATGATGATGACGACCCGACGACGACAACTGCTCAGGCGTCGCG 271
DB 12489 CCGCGCGGTGACACCCCGACCTCGGCGCTGCGCGCTGACACCTCGGCGCGCCCT 12430
QY 272 ACATGACGCGCCATGATTCGCAAGGCGTCAACGAGTGTGACGCGGCAATTCGCGCATCA 331
DB 12429 GCGCGCGCGACGATGCGGAGGATCGCGACGACGCGCTGCGGCTGATCTTCACG 12370
QY 332 GCGTGGCGCGCTGTGGCGACACCCGCGCGCGCGCGCGCTGAGCTGTGAGCAAGGCG 391
DB 12369 CCGCGGACGCGCTTGGCGTGGCGGTGACGCGCGCGCGCGCGCGCTGCGCGCGCG 12310
QY 392 GCTTTACGCTTTCGAGCGCTTTCGACCTGACGCGCTTTCGCGCGCGCGCGCGCG 451
DB 12309 AGGCTTTCAGCTTCA---CGCACCAAGGCGCTTTCAGGCGCGCGCGCGCGCG 12253
QY 452 CCGTCAACGCGCGCTGTGATGATGCGCATTCACGCTGCGCGCGCGCGCGCTGATCGCGACT 511
DB 12252 TCACCGGACGCGCGCTGATGCGCGCGCGCGCGCGCGCGCGCTTCACACTTCGACTTCGAC 12193
QY 512 TGCAGGCG 571
DB 12192 TGCAGGCG 12133
QY 572 TGGCGACG 631
DB 12132 TGGCGCTCACTCTTCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12073
QY 632 CCACCGAAGATCATGAGGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 691
DB 12072 CCGCGTACCG 12013
QY 692 CCCACATGCG 751
DB 12012 AGGCGCTCAACACGACCACTAGTGTGATGATGATGATGATGATGATGATGATGATG 11953
QY 752 GCGCGAGCTGAGAGTGTGCG 811
DB 11952 ACCGCGACTGTGATGAGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11893
QY 812 TGGCG 862
DB 11892 CCGCGGCGCTGCAAGAGCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 11833
QY 863 AGGAGCTGTGCTGAGCG 922
DB 11832 CCGAAGCGCTGTGCG 11773
QY 923 GCGTGTGCTGCG 982

DB 11772 ACATCCGCGCGGTGCG 11713
QY 983 GCGCGGCG 1015
DB 11712 CCG 11680

Search completed: January 14, 2003, 05:58:19
Job time: 1874 secs

```

CDS      complement(4106..5638)
        /gene="CC3449"
        /note="identified by match to protein family HMM"
        /codon_start=1
        /transl_table=11
        /product="ATP synthase F1, alpha subunit"
        /protein_id="AAK25411.1"
        /db_xref="GI:13425167"
        /translation="MDIRAAEISALILKSQIANFCEEAASVDVGVLSVGDGIARIYGL
        DNVAEGMELEPKAGVGMALNTERDNGVAITFGODEIKEGDEVRRIGELVDVPGR
        GLGRVAVPLEPIDGKPIQYTERRRYDVAFCIIPKSVHEPVGTGKSIDPLIPY
        GRGRELHIGRGTGTAVADITLINOKNAGDESAKLYCVVVALIGOKRSTAOIY
        KTLREHGLETTYYVASASAPLQIYIAAPGACMGEMFPDNLHLITLIDDSKOA
        VAIROMSLILRRPGRKATPGDVFYLHRLERAKLNEDGSSLTALPIETQAND
        VSAIPIYVVISITDQIPLFETDLYQGIKRPVNVGIIYSYRGSASQIKAMQVAPIK
        GELAOYREMAAFAGSDLDASTOKMLARGLELTKQPOYAPLSEGVYCIYAGT
        RGYLDGIPTSVSRREAEFLARLHSHQDLLEGIRTKALDKDLENTLKSALDESSST
        FA"
        complement(5691..6245)
        /gene="CC3450"
        /complement(5691..6245)
        /gene="CC3450"
        /note="identified by match to protein family HMM"
        /codon_start=1
        /transl_table=11
        /product="ATP synthase F1, delta subunit"
        /protein_id="AAK25412.1"
        /db_xref="GI:13425168"
        /translation="MADETKATDAGORYAOSLFTLTENGSLQKVEADTKSLKAVAD
        SADLRILASPAFSAEDKGLITAVAKAGGQPLTTKGLIGVANGAVPGLKISAP
        VEISAKHGYTAEVYSAALSPAQLKQVOTALADALCKTEVSTRVPSILGGLKVR
        VGRLEFDSLSKSDLSKFLKRA"
        6338..6742
        /gene="CC3451"
        6338..6742
        /gene="CC3451"
        /note="identified by Glimmer2; putative"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"
        /protein_id="AAK25413.1"
        /db_xref="GI:13425169"
        /translation="MDGRKRAAPSEARILITRAFRRNRONGATPRSLRPSDGISPR
        PRPAPRRRGAHVHARRSGSMWRVPPRCRGRQDRPRHRSRRARRRRRLRRCRE
        RLPVGRKSGCFWSTSPVAGKEKEAAGGPAP"
        complement(6831..7283)
        /gene="CC3452"
        complement(6831..7283)
        /gene="CC3452"
        complement(6831..7283)
        /note="identified by Glimmer2; putative"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"
        /protein_id="AAK25414.1"
        /db_xref="GI:13425170"
        /translation="MLSSDPKRMGICRRRRHPRGRAASTRISFILDLTATSPKPEGLM
        KLRSLAAVALSVSMFAGSAPFADRIIAALDAPVAAKRVVAGGAVFCDETECVSAQ
        APSRALTAACRALAKEVRVAAGFGETKSIDAEDLARCNSAKVAAR"
        7328..7864
        /gene="CC3453"
        7328..7864
        /gene="CC3453"
        /note="identified by Glimmer2; putative"
        /codon_start=1
        /transl_table=11
        /product="hypothetical protein"
        /protein_id="AAK25415.1"
        /db_xref="GI:13425171"
        /translation="WRFASGTADISPPAARPGNRRALKPODLVDPALGLGVDGALF
        DHALGVAVALADPALDVAVGQNGSRKGGIEGAGHOGGHRADGGRDPDRAPPR
        AEDGLQITDOKLAVRLLHHPDHPEDGALDNRGQOORHONHONTRSDKOSERCADRE
        GSHGHAASVTFPFSRTG"
        7951..9258
        gene

```

```

CDS      /gene="CC3454"
        7951..9258
        /gene="CC3454"
        /note="identified by match to PFAM protein family HMM
        PF00441"
        /codon_start=1
        /transl_table=11
        /product="acyl-CoA dehydrogenase family protein"
        /protein_id="AAK25416.1"
        /db_xref="GI:13425172"

Query Match      5.68; Score 99.2; DB 1; Length 14710;
Best Local Similarity 55.58; Pred. 0.014;
Matches 213; Conservative 0; Mismatches 168; Indels 3; Gaps 1;

OY 56 CCTTGACCTGCTGCTCCGGGGGACACCCATCGACGAGCAACACCGGGGGCGC 115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14609 CCTACGATCTGCTCTCGGAGACGGCGGTGTTGTGACGGCGGGCGGACGCTTCC 14550
OY 116 GCGCCGACCTGGGCGTGGCGGCGACCGCATCGCCGCCATCTGTGGACCGCG 175
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14549 AGCCGATGTCGCGTGGCGGCGCATCGTCGCGTGGCGGCAAGGGCTCGCCGCG 14490
OY 176 CCGCGCACACCGGGTGCACGCTGTGGCGCTGTGTGTCGGCCCGGCTTACATCGC 235
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14489 GCGCGAGAGA--GATCGACCGCGCGGCGCATCGTACCCCGGCTTCGATATTC 14433
OY 236 ACACCCAGCAGCAACTACCTGCTCAGCGCGCGACATGACGCCAAGATCTCGCAGG 295
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14432 ACACCCATCATACCGGCCAGCGACCTGAGACGGTACCTTGCCCGCAGCTCGGCCACG 14373
OY 236 GCGTCACACGCGTGTCAAGGGCAATGCGGCATGACGCTGGCGCGTGGCGCACGCCA 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 14372 GCGTACACACGGTGTGTGATGAGGCAATTCGCGCTGCGCGCGTGGCGCGAGC 14313
OY 356 ACCCGCCCGCCCGCCCGCGACGCTGTGAGCAAGGGCGCTTACGTTACAGCGCTTCG 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 14312 ATCGACAGCGCTGATCGCGCTGATGGAAGGCGTGGAGACATCCCTTCCGCTGCA 14253
OY 416 CCGACTACCTGGACCGCGTGGCGG 439
    |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
DB 14252 CCGAGGCGCTGCCCGGGCGTGG 14229

RESULT 15
AX211705 65140 bp DNA linear PAT 06-SEP-2001
LOCUS AX211705
DEFINITION Sequence 1 from Patent W00159126.
ACCESSION AX211705
VERSION AX211705.1 GI:15523937
KEYWORDS
SOURCE
ORGANISM Streptomyces noursei.
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
REFERENCE 1 (bases 1 to 65140)
AUTHORS Zotchev,S.B., Sekurova,O.N., Fjærviik,E., Brautaset,T.,
        Stroem,A.R., Valla,S., Ellingsen,T.E., Sletta,H.V. and
        Guliksen,O.M.
TITLE Gene cluster encoding a nystatin polyketide synthase and its
        manipulation and utility
JOURNAL Patent: WO 0159126-A 1 16-AUG-2001;
        Norges Teknisk Naturvitenskapelige Universitet (NO) ; STIFTTELSEN
        IND OG TEKNISK FORSKNING VED NORGES TEKNISKE HOGSKOLE (NO) ;
        ALPHARMA AS (NO) ; SINVENT AS (NO) ; ZOTCHEV, Sergey Borisovich
        (NO) ; SEKUROVA, Olga Nikolayivna (NO) ; FJÆRVIIK, Espen (NO) ;
        BRAUTASET, Trygve (NO) ; STROM, Arne Reidar (NO) ; VALLA, Svein
        (NO)
FEATURES
        source          Location/Qualifiers
                        1..65140
                        /organism="Streptomyces noursei"
                        /db_xref="taxon:1971"
                        /note="ATCC 11455"

```

BASE COUNT 8270 a 25171 c 22273 g 9426 t
ORIGIN

Query Match 5.6%; Score 99.2; DB 6; Length 65140;
Best Local Similarity 43.5%; Pred. No. 0.0082;
Matches 724; Conservative 0; Mismatches 913; Indels 27; Gaps 5;

QY 62 ACCTGCTGCTGCGGCGGCAACCTTCATGAGCGGCAACACCCCGGGGGCGCG 121
DB 19698 ACGAGCGCTGCGGCGGCACTGCGCTTCAGATGAGCGCGCGCTTCAGGCGCGCG 19757
QY 122 ACCTGCGGCTGCGGCGGCAACCGCATGCGCGCATGCGCGATGCGCGCGCGCG 181
DB 19758 CCGGCGCGGCGGCAACCGAGGCTTCAGCGGAGTGGCGCGCGCGCGGAGACCG 19817
QY 182 ACACCGGCGGAGCTGTCGCGGCGCTGCTGCTGCGCGCGCGCTTCATGCACTGCAACCC 241
DB 19818 ACCCGCGCGCTTCGAGACTGCAACCGCGCGCTGCTGAGCGCGCGCAACACCGCGCGCT 19877
QY 242 ACGAGCACTACCTGCTGCTGAGCGCTGCGACATGACGCGCAAGATCTGCGAGGCGCTCA 301
DB 19878 ACGCGGAGCTGCGGCGGCAATGAGCGCGCGCGCGCTGCGCGGAGAGCGCTGCGC 19937
QY 302 CCAGCGTGTGACGCGGCAATGCGCATGACGCTGCGCGCGCGCTGCGCGCAACCGCGC 361
DB 19938 TCGCGCGCGCGCGCGCGCAACCGCTGCGCGCGCGCGCGCGCGCGCGCGAGACACCG 19997
QY 362 CCGCGCGCGCTGAGCTGCTGAGCAAGCGCGCTTCATCGTTGAGGCGCTTCGCGCACT 421
DB 19998 TCACGATGCGCTGCTGAGCAAGCGCGCGCGCGCGCGCGCTTCGCTGAGCTTCGCTGCT 20057
QY 422 ACCTGAGCGCTTGGCGGCGCAACCGCGCGCGCTTCATCGCGCTGATGCTGCGCGCAAT 481
DB 20058 CCGCGCGCGCTGCGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 20117
QY 482 CAAGCGTGTGCGCGCGCTGCTGAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 541
DB 20118 TCAGATGAGTGAACCGCGCTTCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 20177
QY 542 CGGCGATGCGGAGCTGCGCGGAGGAGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCG 601
DB 20178 TCGTGTGCG 20237
QY 602 CGCGCTTCTACCG 661
DB 20238 CCG 20297
QY 662 CGCTGAGCGCGCGCATGCG 721
DB 20298 TCACGACCTTCACGACCG 20357
QY 722 TGGCGCGCGCTGAGGAAACCTTCGCGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 781
DB 20358 CCGCGCTGCTGCG 20417
QY 782 CGGACGACAGCGTATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 841
DB 20418 TGTGCTGCTGATGCG 20477
QY 842 TCGAGGCGCGCGCATGCG 901
DB 20478 GCGTGTGCTGCG 20537
QY 902 CCACCATGCTCAAGAGGAGCG 961
DB 20538 CCGAGACCG 20591
QY 962 AGCCCTTCCCGAACTGAGCG 1021
DB 20592 ACGAGAGCG 20651
QY 1022 CCAAGTACGAGTGTGCG 1081

DB 20652 TCCGCTGCG 20711
QY 1082 AACCGAGCTGACAGCGCATCTGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1141
DB 20712 GCACCG 20768
QY 1142 CGGACGAGCG 1201
DB 20769 GCCACTGCTGCTGCG 20828
QY 1202 ATGCG 1261
DB 20829 CCGAGCTGACCG 20888
QY 1262 CCG 1321
DB 20889 CCG 20939
QY 1322 TGGTGTGCTGCG 1381
DB 20940 TCGTGTGCTGCG 20999
QY 1382 CG 1441
DB 21000 GCGTGTGCTGCG 21059
QY 1442 TCACCG 1501
DB 21060 GCGACTGCTGAGCG 21119
QY 1502 CCTTAC-----AATCGCGCGTGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1555
DB 21120 CCGGACAGCG 21179
QY 1556 CAACCGCTACATGCG 1615
DB 21180 CCG 21239
QY 1616 AGAAGTGTGCG 1672
DB 21240 TGACAGCG 21299
QY 1673 AAGACTGTGCG 1716
DB 21300 TGACCTTGGAACAGGCG 21343

Search completed: January 14, 2003, 08:07:30
Job time : 5425 secs